28-я Международная Конференция

DSPA-2026

Доклады

28th International Conference
Digital Signal Processing
and its Applications



СЕРИЯ:
ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ЕЁ ПРИМЕНЕНИЕ (выпуск XXVIII)

**Российское
научно-техническое общество
радиотехники, электроники и связи
им. А.С. Попова**

Москва - 2026 г.



XXVIII Международная конференция

Цифровая обработка сигналов

и ее применение

DSPA - 2026

25 марта - 27 марта

AUTEX Ltd.

XXVIII International Conference

Digital Signal Processing

and Its Applications

DSPA - 2026

March 25 - March 27

ПРОГРАММА КОНФЕРЕНЦИИ
CONFERENCE PROGRAMME

Москва - 2026

28-я Международная конференция «Цифровая обработка сигналов и её применение – DSPA-2026», Москва, Россия, доклады.

Серия: Цифровая обработка сигналов и её применение

Выпуск: XXVIII

Рецензенты:

доц. Бабаян П.В., проф. Бартенев В.Г., проф. Брюханов Ю.А., д.т.н. Вашкевич М.И., проф. Витязев В.В., доц. Витязев С.В., проф. Гаврилов К.Ю., prof. Vikram Gadre, доц. Гвоздарёв А.С., член-корр. РАН. Дворкович А.В., д.т.н. Джиган В.И., д.т.н. Зайцев Г.В., к.т.н. Клеопин А.В., д.т.н. Кошелев В.И., доц. Лобов Е.М., д.т.н. Малай И.М., к.т.н. Овинников А.А., доц. Петровский Н.А., д.т.н. Приоров А.Л., проф. Прохоров Ю.Н., проф. Санников В.Г., проф. Сергиенко А.Б., PhD Koushlendra Singh, проф. Сперанский В.С., к.т.н. Стешенко В.Б., prof. Radovan Stojanovic, доц. Стротов В.В., prof. Rui Fan, проф. Федоров В.Ф., к.т.н. Ханян Г.С., доц. Хрящев В.В., проф. Цимбал В.А., проф. Чиров Д.С.

© Авторы докладов
© РНТОРЭС имени А.С. Попова

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи имени А.С. Попова
- Институт радиотехники и электроники имени В.А. Котельникова РАН
- Институт проблем управления имени В.А. Трапезникова РАН
- Компания AUTEX Ltd.

ПРИ УЧАСТИИ:

- Министерство высшего образования и науки РФ
- Федеральное агентство по промышленности РФ
- Институт проблем информатики РАН
- ОАО «РТИ»
- ОАО «ГСКБ «АЛМАЗ-АНТЕЙ»
- ЗАО «Инструментальные системы»
- НТЦ «МОДУЛЬ»
- ОАО «Концерн радиостроения «Вега»
- ОАО «Концерн «СОЗВЕЗДИЕ»
- ГУП НПЦ «Элвис»
- Владимирский государственный университет
- Всероссийский научно-исследовательский институт физико-технических и радиотехнических измерений «ВНИИФТРИ»
- Московский авиационный институт
- Московский государственный технический университет им. Н.Э. Баумана
- Московский технический университет связи и информатики
- Московский Физико-технический институт (университет)
- Московский технологический университет
- Московский институт электронной техники
- Московский энергетический институт (университет)
- Рязанский государственный радиотехнический университет
- Санкт-Петербургский государственный университет телекоммуникаций имени проф. М. А. Бонч-Бруевича
- Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
- Ульяновский государственный технический университет
- Ярославский государственный университет

МЕЖДУНАРОДНЫЙ ПРОГРАММНЫЙ КОМИТЕТ КОНФЕРЕНЦИИ:

Председатель Международного программного комитета:
академик РАН **Гуляев Юрий Васильевич**

Сопредседатели Международного организационного комитета: академик РАН **Васильев С.Н.**, академик РАН **Каляев И.А.**, **Кирпичников А.П.**, академик РАН **Сигов А.С.**

Председатель Программной группы Международного организационного комитета:
член-корр. РАН **Дворкович Александр Викторович**

Заместитель Председателя Программной группы Международного организационного комитета:
профессор **Витязев Владимир Викторович**

Программная группа Международного организационного комитета:
проф. Аджемов А.С. (Россия), член-корр. РАН Борисов В.И. (Россия), д.т.н. Борисов Ю.И. (Россия), проф. Викрам Гадре (Индия), проф. Михалис Зервакис (Греция), проф. Кушлендра Сингх (Индия), академик РАН Никитов С. А. (Россия), академик РАН Новиков Д.А. (Россия), проф. Радован Стоянович (Черногория), проф. Жуй Фань (КНР), академик РАН Чаплыгин Ю.А. (Россия)

Административная группа Международного организационного комитета:
доц. Бабаян П.В., проф. Бартенев В.Г., проф. Брюханов Ю.А., проф. Васильев К.К., доц. Витязев С.В., проф. Гаврилов К.Ю., проф. Джиган В.И., Зубарев Р.В., к.т.н. Клеопин А.В., доц. Лобов Е.М., д.т.н. Малай И.М., д.т.н. Приоров А. Л., проф. Прохоров Ю.Н., проф. Санников В.Г., доц. Сергиенко А.Б., проф. Сперанский В.С., доц. Стешенко В.Б., проф. Федоров В.Ф., доц. Хрящев В.В., проф. Чиров Д.С.

ПЛЕНАРНЫЕ ДОКЛАДЫ **1**

Секция 1. ТЕОРИЯ СИГНАЛОВ И СИСТЕМ **7**

Секция 2. ТЕОРИЯ И МЕТОДЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ **65**

Секция 3. ОБРАБОТКА СИГНАЛОВ В СИСТЕМАХ ТЕЛЕКОММУНИКАЦИЙ . . **91**

Секция 4. ОБРАБОТКА СИГНАЛОВ В РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ . . . **112**

Секция 5. ОБРАБОТКА И ПЕРЕДАЧА ИЗОБРАЖЕНИЙ **127**

Секция 6. ОБРАБОТКА СИГНАЛОВ В БИОТЕХНИЧЕСКИХ СИСТЕМАХ **145**

Секция 7. ПРОЕКТИРОВАНИЕ И ТЕХНИЧЕСКАЯ РЕАЛИЗАЦИЯ СИСТЕМ ЦОС 161

Секция 8. СИСТЕМЫ «ТЕХНИЧЕСКОГО ЗРЕНИЯ», ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, РОБОТОТЕХНИКА **181**

Секция 9. ОБРАБОТКА СИГНАЛОВ В ИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ **246**



Галяев Андрей Алексеевич
чл.-корр. РАН, Институт проблем управления им. В.А. Трапезникова РАН, Москва

Оптимизационные постановки в задачах обнаружения

В докладе приведены оптимизационные постановки, возникающие при формализации задачи обнаружения детерминированного сигнала при наличии шума среды, которая, на первый взгляд, является задачей бинарной классификации. Однако грамотная декомпозиция задачи поднимает вопросы о том, какая информация о свойствах сигнала и шума известна и как правильно понимать решение этой задачи.

В общем случае задача обнаружения представляет собой бинарную классификацию, для которой требуется сформировать правило принятия решения о наличии или отсутствии полезного сигнала в сигнально-шумовой смеси. При этом, в случае многоклассовой классификации, признаки, сформированные на основе входящей временной последовательности отсчетов, могут относится к множеству различных классов, на которые делится как полезный сигнал, так и шум. Чем больше классов предлагается для принятия решения, тем точнее требуется знать частотно-временные свойства сигнала и шума. В бинарной классификации различение двух гипотез опирается на функцию различения, которая показывает предельные возможности метода обнаружения детерминированного сигнала в шуме. Это можно рассматривать как некоторый аналог принципа неопределенности в квантовой механике или цифровой обработке сигналов (Теорема Котельникова), поскольку, даже увеличивая количество признаков нельзя решить задачу обнаружения сигнала в шуме при малом отношении сигнал/помеха. Характеристики сигнала можно отнести к локальным или интегральным информационным характеристикам, которые являются функциями (функционалами) от входящего временного ряда. Поэтому требуется сформировать оптимальные правила для синтеза информационной системы, решающей задачу обнаружения на основе ограниченной априорной информации. Однако стоит учитывать, что особенностью такой оптимальной обработки является неустойчивость к изменению статистических свойств шума и априорная неопределенность о свойствах сигнала, и тот факт, что многие математические конструкции оказываются не реализуемыми или реализуемыми с большими погрешностями.

Текущий доклад посвящен обсуждению свойств задачи обнаружения и формированию подхода для улучшения качества решения. Будут рассмотрены несколько последовательных этапов формализации и решения задачи обнаружения, опираясь на известные формулировки задачи разделения двух гипотез и Леммы Неймана-Пирсона. Затем будут выделены критерии, свойства и признаки детерминированного сигнала, применены методы оптимальной обработки сигналов с вычислением вероятностных временных, частотных и частотно-временных распределений, и на финальном этапе построены статистики для модельных экспериментов.

Andrey Alekseevich Galyaev
Cor. Mem. of RAS, V.A. Trapeznikov Institute of Control Problems of the Russian Academy of Sciences, Moscow

Optimization statements in detection tasks

The report presents optimization statements that arise when formalizing the problem of detecting a deterministic signal in the presence of ambient noise, which at first glance is a binary classification problem. However, decomposing the problem raises questions about what information about the properties of the signal and noise is known and how to properly understand the solution to this problem.

In general, the detection task is a binary classification, for which it is necessary to form a rule for deciding on the presence or absence of a useful signal in the signal-noise mixture. In this case, the signs formed on the basis of the incoming time sequence of samples can belong to many different classes, into which both the useful signal and the noise are divided. The more classes are offered for decision-making, the more precisely it is required to know the time-frequency properties of the signal and noise. In binary classification, the distinction between two hypotheses is based on the discrimination function, which shows the limiting capabilities of the deterministic signal detection method in noise. This can be considered as some kind of analogue of the uncertainty principle in quantum mechanics or digital signal processing (Kotelnikov's Theorem), since even increasing the number of features cannot solve the problem of detecting a signal in noise with a small signal-to-noise ratio. It turns out that the key physical characteristics of the signal-noise mixture that are known or can be measured are the type and length of the time window, sampling frequency, signal-to-noise ratio, statistical properties of noise, and the amount of discretionary signal.

The remaining characteristics are attributed to local or integral information characteristics, which are functions (functionals) of the incoming time series. Therefore, it is necessary to form optimal rules for the synthesis of an information system that solves the problem of detection based on limited a priori information. However, it should be borne in mind that the peculiarity of such optimal processing is the instability to changes in the statistical properties of noise and a priori uncertainty about the properties of the signal, and the fact that many mathematical constructions turn out to be unrealizable or implementable with large errors.

The current report is devoted to discussing the properties of the detection problem and developing an approach to improve the quality of the solution. Several successive stages of formalization and solution of the detection problem will be considered, based on the well-known formulations of the problem of separation of two hypotheses and the Neiman-Pearson Lemma. Then the criteria, properties and features of a deterministic signal will be highlighted, optimal signal processing methods will be applied with the calculation of probabilistic time, frequency and time-frequency distributions, and at the final stage statistics for model experiments will be constructed.



Усатюк Василий Станиславович
руководитель группы теории информации компании ООО «Т8», Москва, научный сотрудник Юго-Западного государственного университета, кандидат технических наук

Графовое сжатое измерение и вложения признаков глубоких нейронных сетей

Разреженное преобразование Фурье и теория сжатого измерения (Compressed Sensing, CS) произвели революцию в магнитно-резонансной томографии, позволив радикально снизить энергопотребление и вычислительную сложность. Сегодня эти методы широко применяются для восстановления сигналов в таких областях, как медицинская визуализация, квантовая томография, радиолокация и связь. Ключевым объектом рассмотрения в данном докладе является свойство ограниченной изометрии на графе (Graph Restricted Isometry Property, RIP-2). Приводится интуитивно понятная интерпретация этого свойства с точки зрения обработки сигналов, которая гарантирует устойчивое и помехозащищенное восстановление сигнала. Также рассматриваются последние достижения — от ускоренных алгоритмов разреженного быстрого преобразования Фурье до сжатия информации о состоянии канала в системах беспроводной связи, соответствующих 18-й редакции стандарта 5G.

Далее излагается современный подход, объединяющий сжатое измерение на графах с разреженными вложениями пространства признаков глубоких нейронных сетей. Классические задачи машинного обучения, такие как кластеризация (обучение без учителя), классификация и регрессия (обучение с учителем), с позиции сигнальной обработки могут быть сформулированы как задачи восстановления нелинейных сигналов разреженных вложений (embeddings) признаков на графах. При этом глубокие сети выступают в роли преобразований, осуществляющих кластеризацию признаков. Графовая разреженная обработка нелинейных признаков выполняет регуляризацию, улучшая обусловленность матриц и сокращая необходимое количество признаков.

Практическое влияние подхода иллюстрируется двумя сценариями развертывания. Граничный ИИ (Edge-AI): устройства со сверхнизким энергопотреблением, расположенные рядом с источником сигнала, которые выполняют логический вывод (inference) в реальном времени на основе значительно меньшего числа измерений. Центры обработки данных (ЦОД): высокопроизводительные системы, использующие структурную разреженность на графах для ускорения обработки больших пакетов данных (сигналов, изображений или текстов) без потери точности.

Таким образом, объединяя теорию сжатого измерения (CS), обработку сигналов на графах и глубокое обучение, данный подход обеспечивает универсальность принципов сжатого измерения в нелинейной теории сигнальной обработки — экономя энергию на периферийных устройствах и повышая пропускную способность за счет масштабного параллелизма при обработке данных в ЦОД.

Усатюк Василий Станиславович, руководитель группы теории информации компании ООО «Т8», Москва, научный сотрудник Юго-Западного государственного университета, кандидат технических наук. Он работает в областях теории информации, цифровой обработки сигналов и машинного обучения. Его исследования сосредоточены на разработке алгоритмов и программно-аппаратных архитектур для глубоких нейронных сетей, помехоустойчивого кодирования (FEC), а также передовых решений для квантовых вычислений, сенсорных систем и систем связи. Ранее он занимал должность старшего инженера-исследователя в ООО "Техкомпании Хуавей" и руководил лабораториями "Параллельные вычисления" и "Технические средства защиты информации" в Братском государственном университете. Василий является автором более 45 международных патентов и внес вклад в отраслевые стандарты в области постквантовой криптографии и помехоустойчивого кодирования для сетей 5G.

Vasiliy Stanislavovich Usatyuk
Head of the Information Theory Group at T8 Ltd., Moscow, researcher at South-West State University, candidate of sciences

Graph-Based Compressed Sensing Meets Deep Neural Network Feature Embedding's

The Sparse Fourier Transform and compressed sensing (CS) have reshaped magnetic resonance imaging—drastically reducing power consumption and computational load—and now enable high fidelity reconstruction across biomedical imaging, radar, synthetic aperture satellite systems, communications, remote sensing, and quantum tomography. This talk centers on the second Restricted Isometry Property, providing an intuitive signal processing interpretation of its guarantee for stable, noise robust recovery and surveying recent advances —from accelerated sparse FFT algorithms to massive MIMO channel state information compression for 5G Release 18. A modern pipeline that merges graph based CS with deep neural sparsifiers is then introduced. Classical machine-learning tasks (clustering or classification) are recast as learning sparse embedding on graphs, where deep networks act as data driven sparsifiers that precondition signals before measurement, improving matrix conditioning and cutting required samples. Two deployment scenarios illustrate the impact. Edge AI scenarios – ultra low power devices performing real time inference with far fewer measurements. Data center scenarios – high throughput systems leveraging graph structured sparsity to accelerate large batch (images or text tokens) processing without accuracy loss. By uniting CS theory, graph signal processing, and deep learning, the approach delivers universally efficient sensing—saving power on the edge while boosting throughput at scale.

Vasiliy Stanislavovich Usatyuk is the Head of the Information Theory Group at T8 Ltd. in Moscow and a researcher at South-West State University. He works in the fields of information theory, digital signal processing, and machine learning. His research focuses on developing algorithms and CPU, GPU, QPU, FPGA and ASIC oriented architectures for deep neural networks, forward error correction, and advanced solutions for quantum computation, sensing, and communication systems. He has previously held a senior research engineer position at Huawei and led the Parallel Computation and Information Security Laboratory at Bratsk State University. Vasiliy is the author of more than 45 international patents and has contributed to industry standards in post-quantum cryptography and 5G error correction



Волохов Владимир Андреевич
старший научный сотрудник ООО «ЦРТ-инновации» (г. Санкт-Петербург), доцент Национального исследовательского университета ИТМО, к.т.н., доцент.

Современные тенденции развития голосовой биометрии

В рамках доклада рассматривается конвейер современной системы голосовой биометрии с акцентом на устройство блока построения дикторских моделей (дикторских эмбеддингов), в основе которого лежат глубокие нейросетевые архитектуры такие, как свёрточные и трансформерные. Приводятся сведения по обучению нейронных сетей применительно к задаче распознавания диктора, а также оценки количества параметров, используемых в них, и временной сложности, позволяющие понять практическую применимость подобных алгоритмов. Рассматривается оценка качества современных систем распознавания диктора в рамках типовых бенчмарков, а также обсуждаются перспективные направления развития голосовой биометрии..

Волохов Владимир Андреевич, старший научный сотрудник ООО «ЦРТ-инновации» (г. Санкт-Петербург), к.т.н., доцент Национального исследовательского университета ИТМО. Занимается практической деятельностью в области цифровой обработки речевых сигналов: распознавание диктора и разговорного языка, дикторская и языковая диаризация. Выступает руководителем учебной дисциплины «Распознавание диктора» в Национальном исследовательском университете ИТМО. Участник международных конференций SPECOM, Odyssey, INTERSPEECH и ICASSP.

Vladimir Andreevich Volokhov
Senior researcher at STC-innovations Ltd., Saint Petersburg, associate professor of ITMO University, candidate of sciences, associate professor

Current Trends in the Development of Voice Biometrics

The report discusses the pipeline of a modern voice biometrics system. Particular attention is paid to the description of the block for computing of speaker embeddings, which is based on deep neural network architectures such as CNN and Transformers. Some information is provided on training these neural networks in relation to the speaker recognition task, as well as estimates of the number of parameters used and the time complexity, which make it possible to understand the practical applicability of such algorithms. The quality assessment of modern speaker recognition systems on standard benchmarks and promising directions for the development of voice biometrics are considered.

Vladimir Andreevich Volokhov – senior researcher at STC-innovations Ltd., Saint Petersburg, candidate of sciences, associate professor of ITMO University. He works in the field of digital speech signal processing: speaker and spoken language recognition, speaker and language diarization. He is the

head of the Speaker Recognition course at ITMO University and is a participant in international conferences such as SPECOM, Odyssey, INTERSPEECH and ICASSP

НЕЙРОСЕТЕВЫЕ АЛГОРИТМЫ ОЦЕНКИ АМПЛИТУДЫ РАДИОСИГНАЛА С НЕИЗВЕСТНЫМИ ДЛИТЕЛЬНОСТЬЮ, ЧАСТОТОЙ И НАЧАЛЬНОЙ ФАЗОЙ

доц. Корчагин Ю.Э., асп. Нгуен Ван Тхю

Воронежский государственный университет

Исследованы алгоритмы оценки амплитуды ускополосного радиосигнала произвольной формы с неизвестными частотой, длительностью и начальной фазой, наблюдаемого на фоне аддитивного гауссовского белого шума. Использован квазиправдоподобный метод для преодоления априорной параметрической неопределённости, где вместо неизвестных параметров используются их ожидаемые значения. Построены нейросетевые алгоритмы оценки, а также выполнен анализ построенного алгоритма для разных типов входных данных нейросети.

Введение

Задача оценки амплитуды радиосигнала с неизвестными параметрами давно известна научной и инженерной общественности и неоднократно рассматривалась в литературе. Так, например, в работах [1,2,3] рассмотрены алгоритмы оценки амплитуды с неизвестными длительностью и начальной фазой, поскольку они обычно неизвестны на приёмной стороне в силу особенностей распространения сигнала. Однако в ряде практических приложений, частота – из-за доплеровского сдвига и рассинхронизации опорных генераторов приёмника и передатчика также оказывается неизвестной. В задачах оценки амплитуды сигнала при неизвестных длительности, начальной фазе и частоте для преодоления априорной параметрической неопределённости можно пользоваться различными методами. В данной работе предложены нейросетевые алгоритмы и квазиправдоподобный алгоритм обработки таких радиосигналов, а также выполнен сравнительный анализ алгоритмов оценки.

Постановка задачи

Пусть на интервале времени [0,T] доступна наблюдению реализация

$$\xi(t) = s(t, a_0, \varphi_0, \omega_0, \tau_0) + n(t) \quad (1)$$

аддитивной смеси полезного сигнала

$$s(t, a_0, \varphi_0, \omega_0, \tau_0) = \begin{cases} a_0 f(t)\cos(\omega_0 t - \varphi_0), 0 \le t \le \tau_0, \\ 0, \qquad t < 0, t > \tau_0 \end{cases} \quad (2)$$

и гауссовского белого шума $n(t)$ с односторонней спектральной плотностью N_0. где $f(t)$ - непрерывная функция, описывающая форму огибающей, $a_0, \varphi_0, \omega_0, \tau_0$ – неизвестные амплитуда, начальная фаза, частота и длительность принимаемого сигнала соответственно. Полагаем, что длительность сигнала принимает значения из априорного интервала.

$$\tau_0 \in [T_1, T_2].$$

Необходимо сформировать оценку амплитуды полезного сигнала (2), считая длительность, начальную фазу и частоту радиосигнала неизвестными параметрами, в оценке которых нет необходимости. Наряду с общепринятыми статистически оптимальными алгоритмами оценки можно использовать нейросетевые алгоритмы, что особенное актуально, когда часть априорных сведений о сигнале или шуме остаётся неизвестной. В данной работе предложены два нейросетевых алгоритма, входными данными которых являются дискретные отсчёты наблюдаемой реализации или дискретное представление её спектрограммы.

Квазиправдоподобный алгоритм оценки амплитуды

Для оценки амплитуды можно использовать метод максмального правдоподобия (МП) [1,2], согласно которому оценка амплитуды при априори известных длительности, начальной фазе и частоте совпадает с положением абсолютного (наибольшего) максимума логарифма функционала отношения правдоподобия (ФОП):

$$a_{0m} = \underset{a}{argsup} L(a, \varphi_0, \omega_0, \tau_0)$$

При неизвестных длительности, амплитуде, частоте и начальной фазе логарифм ФОП зависит от четытрех неизвестных параметров:

$$L(a, \varphi, \omega, \tau) = \frac{2a}{N_0}\int_0^{\tau} f(t)\cos(\omega t - \varphi)[\xi(t) - af(t)\cos(\omega t - \varphi)/2]dt.$$

Следовательно, имеется априорная параметрическая неопределенность относительно длительности, частоты и начальной фазы. В качестве одного из возможных способов преодоления этой неопределенности можно использовать квазиправдоподобный (КП) алгоритм оценки, согласно которому вместо неизвестных длительности τ, частоты ω и начальной фазы φ в выражении (5) используются их некоторые ожидаемые значения $\tau^*, \omega^*, \varphi^*$ соответственно. Согласно работе [3], при равенстве ожидаемых и истинных значений параметров, КП алгоритм оценки амплитуды обладает смещением и рассеянием:

$$B = 0, \quad V = \frac{a_0^2}{z_0^2}$$

$z_0^2 = a_0^2 q(\tau_0)$ - отношение сигнал/шум (ОСШ) на выходе приёмника МП для сигнала с единичной амплитудой и длительностью τ, а

$$q(\tau) = \frac{1}{N_0}\int_0^\tau f^2(t)dt.$$

Нейросетевые алгоритмы оценки амплитуды радиосигнала

Рассмотрим далее неросетевой алгоритм оценки амплитуды, внутренняя структура которого не связана с оптимальными или квазиоптимальными алгоритмами обработки, а построена по аналогии с нейростевыми конструкциями различения образов на изображениях [4,5].

Применение нейросети предполагает формирование радиосигнала с неизвестными параметрами и белого шума в виде дискретных отсчётов с частотой дискретизации f_s. На интервале наблюдения [0,T] располагается $N = int(Tf_s)$ отсчётов. В процессе формирования принятой реализации с шагом во времени $\Delta t = 1/f_s$ вычислялись отсчеты функций радиосигнала и белого шума. Таким образом, на вход нейросети поступают дискретные отсчёты $\xi_i,\ t_i = i\Delta t, i = \overline{1,N}$. Выходом нейросети должно являться значения амплитуды. Реализация в виде дискретных отсчётов имеет следующий вид:

$$\xi_i = af(t_i)\cos(\omega t_i - \varphi) + n_i.$$

Чтобы достичь качества работы нейросети, близкого к оптимальному, необходимо разработать соответствующую нейросеть и обеспечить достаточно большой объём обучающих данных, чтобы модель могла изучить характеристики сигнала и шума. Данный подход позволяет получить оценку значения истинной амплитуды сигнала

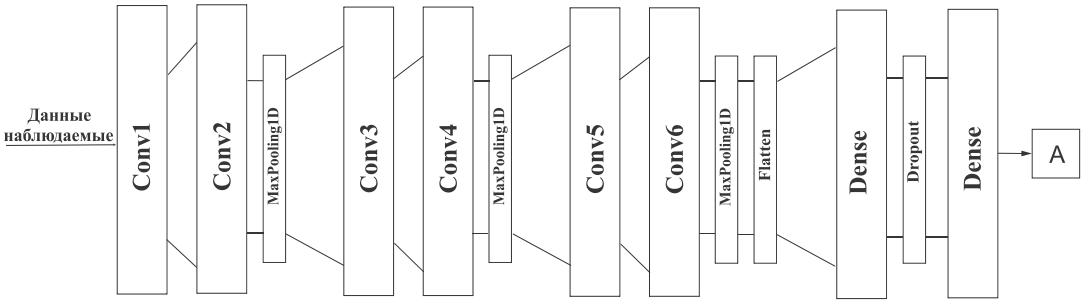


Рисунок 1 – Блок-схема нейросети оценки амплитуды

На рисунке 1 показана блок-схема одной из возможных нейросетей, выбранная в работе. Нейронная сеть построена с использованием архитектуры Sequential для распознавания данных, представленных в виде достаточно длинного вектора. Модель включает три сверточных блока: каждый блок состоит из двух сверточных слоев с активацией ReLU за которыми следует MaxPooling1D для уменьшения размерности признаков. Обучение модели осуществляется с оптимизатором Adam, функцией потерь “mse”, котораясоответствует минимизации среднеквадратической ошибки.

Здесь обозначено: Conv1, Conv2 – свёрточные слои, которые содержат по 16 фильтров с ядрами (длиной импульсной характеристики) 9 отсчётов, Conv3, Conv4 – свёрточные слои, которые содержат по 32 фильтра с длиной импульсной характеристики 9 отсчётов и Conv5, Conv6 – свёрточные слои, которые содержат по 64 фильтра с длиной импульсной характеристики 9 отсчётов. Были выбраны функции активации свёрточных слоёв ReLu.

MaxPooling – слои прореживания отсчётов сигнала в 3 раза, в которых входная последовательность разбивается на неперекрывающиеся окна по 3 элемента, в каждом окне выбирается максимальное значение, в результате получаем новую последовательность, укороченную в 3 раза. После сверточных блоков применяется слой Flatten и первый полносвязный слой Dense с 128 нейронами и ReLU, Dropout 0.5. Второй полносязный слой обладает одним нейроном. В результате на выходе второго полносязного слоя получим значение оценки амплитуды.

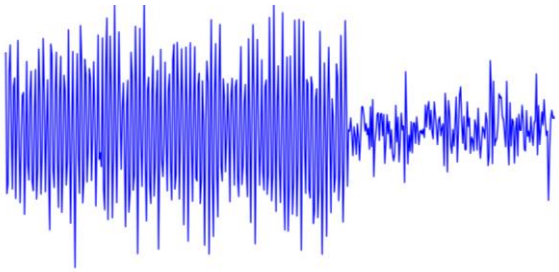


Рисунок 2 – Пример формы принятой реализации

Для обучения нейросети был сгенерирован набор размеченных данных, который содержал отсчёты реализаций, подаваемых на вход сети и поставленное в соответствие каждой реализации истинное значение амплитуды сигнала, которое использовалось как целевое значение оценки. Общий объём обучающей выборки составлял 40000 реализаций. Для каждой реализации отношение сигнал/шум (ОСШ) выбиралось случайным образом из широкого диапазона значений от 0.1 до 100. Такой подход позволяет охватить различные сигнальные и помеховые условия, тем самым повышая обобщающую способность и надёжность модели при определении амплитуды сигнала в широком диапазоне уровней шума. После обучения модели выполнено статистическое моделирование нейросети путём расчёта смещения и рассеяния, допущенных моделью по сравнению с истинными значениями. Данный процесс повторяется 10000 раз, после чего вычисляются средние значения смещения и рассеяния. Амплитуда сигнала выбиралась случайным образом из равномерного распределения в диапазоне от 0.1 до 10 что позволяет модели усваивать амплитудные характеристики сигнала даже при наличии шума.

Вторая нейросетевая модель измерителя амплитуды основана на анализе спектрограммы принятой реализации, пример которой изображен на рисунке 3. При вычислении спектрограммы было использовано окно hamming, с наложением 1/2. Структрура нейросети была выбрана аналогично структуре, показанной на рисунке 1, где в качестве наблюдаемых данных использовались дискретные отсчёты спектрограммы, а свёрточные слои были двумерными (conv2d). Все параметры сигнала, белого шума, а также объём входных обучающих данных полностью соответствуют значениям, использованным в первой нейросетевой модели. Это позволяет осуществить наглядное сравнение эффективности двух моделей обучения.

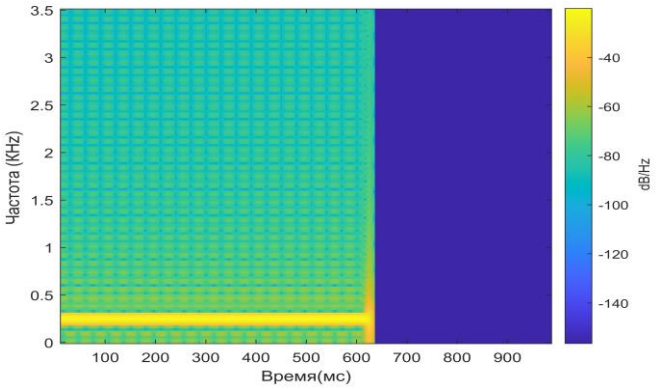


Рисунок 3 – Пример формы спектрограммы принятой реализации

В качестве примера рассмотрим характеристики оценки амплитуды радиосигнала, огибающая которого имеет форму:

$$f(t) = \sqrt{\frac{2b}{1 - e^{-2b}}}\, exp\left[\frac{-bt}{T_2}\right].$$

Введём в рассмотрение величину $z_r^2 = a_0^2 T_2 / N_0$ - ОСШ для прямоугольного радиоимпульса амплитуды и длительности T_2. Выразим ОСШ z_0^2 для принятого сигнала через ОСШ z_r^2

$$z_0^2 = z_r^2 \frac{e^{-2bx_0 - 1}}{e^{-2b-1}},$$

где $x_0 = \tau_0 / T_2$. Формирование реализаций, обучение сети и моделирование её работы выполнялось в нормированном времени $x = t/T_2$, были выбраны значения $b = 1$, $\frac{T_2}{T_1} = 10$. Частота гармонического заполнения выбиралась из диапазона $[f_{min}, f_{max}]$, где максимально возможная частота соответствовала $\kappa = 100$ периодов гармонического заполнения на минимальной длительности сигнала. Частота дискретизации выбиралась как $f_s = 10 f_{max}$. В ходе моделирования многократно генерировалась реализация смеси сигнала с шумом, не принадлежащая обучающему набору. Реализация подавалась на вход нейростети первого типа, а её спектрограмма – на вход нейросети второго типа. Накапливалась ошибка оценки амплитуды, а также квадрат ошибки, на основе которых вычислялись выборочные значения смещений и рассеяния. Было выполнено 10^4 циклов испытаний.

На рисунках 4 и 5 изображены соответственно зависимости нормированных условных рассеяния $V = \langle (A - a_0)^2 \rangle / a_0^2$ и смещения $b = \langle A - a_0 \rangle / a_0$, где A – значение оценки на выходе нейросети. Результаты статистического моделирования нейросетевых алгоритмов показаны на рисунках 4 и 5 квадратиками для исходных данных в виде отсчётов реализации и треугольниками для исходных данных в виде отсчётов спектрограммы. Сплошная линия соответствует зависимости условного рассеяния от ОСШ при отсутствии расстройки параметров.

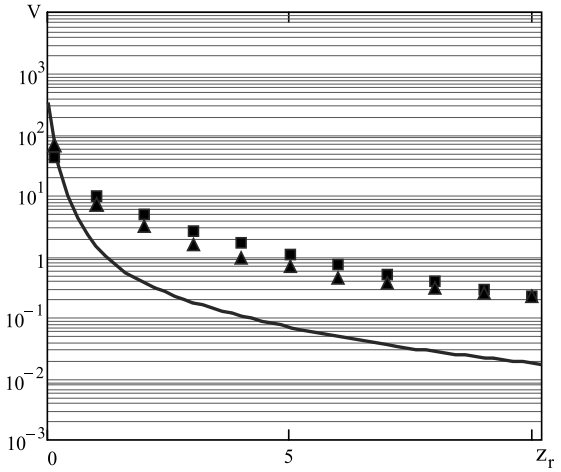


Рисунок 4 – Зависимость условного рассеяния от ОСШ

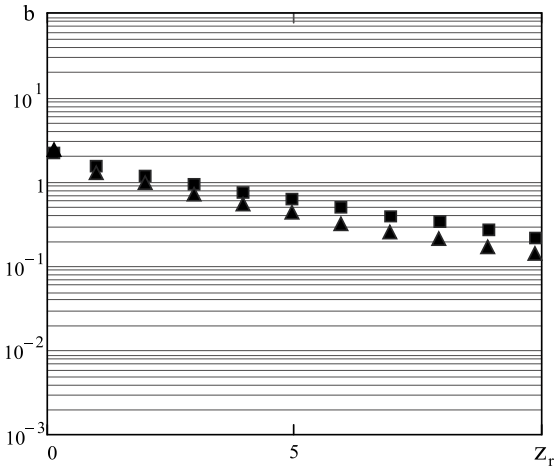


Рисунок 5 – Зависимость условного смещения от ОСШ

Выводы

Нейросетевые алгоритмы оценки амплитуды проигрывают по сравнению со статистически оптимальными. Их применение целесообразно, при невозможности построения оптимальных измерителей ввиду недостатка априорных сведений. При схожей структуре нейросетей лучшей эффективностью функционирования обладает сеть, основанная на анализе спектрограммы.

Литература

1. Е.И. Куликов, А.П. Трифонов. Оценка параметров сигналов на фоне помех // М.: Сов. Радио. 1978. —296с.
2. А.П. Трифонов. Совместное различение сигналов и оценка их параметров на фоне помех // М.: Радио и связь. 1986. — 268с.
3. А.П. Трифонов, Ю.Э. Корчагин, М.В. Трифонов, Е.В. Литвинов. Характеристики оценки амплитуды радиосигнала с неизвестными длительностью и фазой // Радиотехника и электроника. 1991. №2. С. 319–327.
4. С. И. Николенко, А. А. Кадурин, Е. О. Архангельская. Глубокое обучение. Погружение в мир нейронных сетей // СПб. : Питер. 2018. – 480 с.
5. С. Хайкин. Нейронные сети и обучение машин // Пер. с англ. — М. : Вильямс, 2016. — 688 с.

NEURAL NETWORK ALGORITHMS FOR ESTIMATING THE AMPLITUDE OF A RADIO SIGNAL WITH UNKNOWN DURATION, FREQUENCY AND INITIAL PHASE

Korchagin Y.E.[1], Nguyen Van Thuy[1]

[1]Voronezh State University

Algorithms for estimating the amplitude of a narrowband radio signal of arbitrary waveform, observed in the presence of additive white Gaussian noise, with unknown frequency, duration and initial phase are investigated. A quasi-likelihood approach is employed to overcome a priori parametric uncertainty, in which the unknown parameters are replaced by their expected values. A neural network–based estimation algorithm is developed, and a performance analysis of the proposed algorithm is carried out for various types of input data.

АНАЛИЗ ВОЗМОЖНОСТЕЙ ОПРЕДЕЛЕНИЯ КООРДИНАТ БПЛА ПО СИГНАЛАМ БОРТОВОГО МОДЕМА LTE

студ. Рябчицкий М.М., студ. Илларионова П.В.

Московский технический университет связи и информатики

В современных реалиях имеется большая потребность в отслеживание различных беспилотных аппаратов и их траекторий движений, а также в определении и классификации БПЛА среди прочих абонентов сети. В данной статье рассматриваются различные методы определения координат БПЛА по сигналам бортового модема LTE, через базовые станции. Приводится подробное описание каждого из методов и их реализации, их преимущества и недостатки по таким критериям, как: сложность реализации, точность, надёжность и стоимость.

Обнаружение БПЛА является важной задачей, как при решении специальных задач, так и при контроле соблюдения движения безэкипажных воздушных средств (БВС) в условиях «умного города». Типовыми средствами обнаружения и измерения координат БПЛА являются специализированные радиолокационные средства и средства оптического наблюдения. У каждого типа данных средств есть свои преимущества и недостатки. В частности, радиолокационные средства являются всепогодными, но не обладают скрытностью применения. Оптические средства наблюдения являются пассивными и обладают скрытностью применения, но возможность их применения зависит от погодных условий.

Анализ существующих перспективных подходов к обнаружению БПЛА показывает, что в настоящее время появилось большое количество публикаций по обнаружению БПЛА с использованием сторонних сигналов, например сигналов сотовой связи (3/4/5G) или радиовещания (например DBV-T/T2). Данное направление является актуальным.

В сетях мобильной связи стандарта LTE для решения задачи позиционирования абонентов используют методы, основанные на различных методах, основными из которых являются: пространственная обработка сигналов, разностно-дальномерные методы и методы пеленгации, а также алгоритмы, использующие ресурсы наземной инфраструктуры сети (Cell ID, Cell ID-TA, OTD, E-OTD).

Метод Cell ID это самый простой метод позиционирования абонентов который реализован не только в сетях LTE, но и в сетях предыдущих поколений GSM и UMTS. Он предусматривает определение местоположения абонента по активной соте, в которой он находится. Преимуществом метода Cell ID является то, что метод не требует доработки аппаратуры абонента. Погрешность определения координат абонента в данном случае соизмерима с размерами соты (ячейки) и может достигать порядка 150 м в городе и 20-30 км в сельской местности. Такая точность слишком мала, и назвать эти величины приемлемыми нельзя

Другим простым методом позиционирования абонентов в сетях LTE является метод измерения разности времени прихода сигнала (Time difference of arrival (TDoA)). В сетях LTE данный метод может быть реализован на базе сети с помощью параметра Timing advance (TA). Параметр Timing advance используется для синхронизации в канале сети LTE путем компенсации временной задержки времени прохождения сигнала. То есть, TA это параметр времени упреждения ответного сигнала (рисунок 1). Так как данная информация используется для штатной работы сети, то метод позиционирования абонентов с ее помощью не требует дополнительных затрат на модернизацию оборудования.

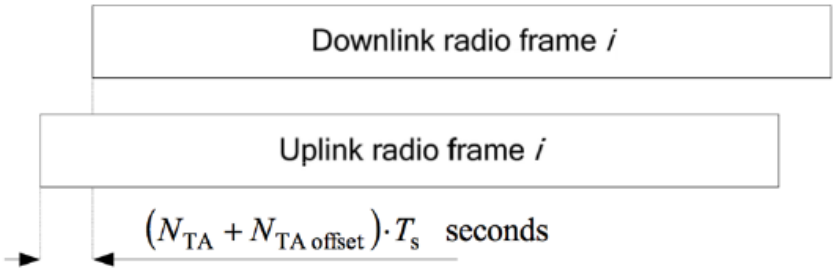


Рисунок 1 – Значение TA на физическом уровне

Метод TDoA применяется, если имеются измерения разности времен распространения сигналов до БС (NodeB – eNB). Элемент сети LTE, Enhanced Serving Mobile Location Centre (E-SMLC), предоставляет пользовательскому оборудованию (User Equipment, UE) перечень соседних базовых станций. Пользовательское оборудование проводит измерения времен распространения сигнала и передает результат на сервер TDoA. Эти измерения преобразуются в разности расстояний Δr_{ij} до точек БС (eNB) с известными координатами.

Так как координаты -й БС x_{ci}, y_{ci} известны, то координаты БПЛА x, y могут быть определены при решении системы нелинейных уравнений с использованием разностно-дальномерного метода по оценке TDoA:

$$\Delta r_{ij} = \sqrt{(x_{ci} - x)^2 + (y_{ci} - y)^2} - \sqrt{\left(x_{cj} - x\right)^2 + \left(y_{cj} - y\right)^2}.$$

Решение может быть записано в виде

$$x = \frac{\left[c_4 - \frac{\Delta r_4}{\Delta r_2} c_2\right]\left[d_3 - \frac{\Delta r_3}{\Delta r_2} d_2\right] - \left[c_3 - \frac{\Delta r_3}{\Delta r_2} c_2\right]\left[d_4 - \frac{\Delta r_4}{\Delta r_2} d_2\right]}{\left[d_3 - \frac{\Delta r_3}{\Delta r_2} d_2\right]\left[b_4 - \frac{\Delta r_4}{\Delta r_2} b_2\right] - \left[b_3 - \frac{\Delta r_3}{\Delta r_2} b_2\right]\left[d_4 - \frac{\Delta r_4}{\Delta r_2} d_2\right]},$$

$$y = \frac{\left[c_3 - \frac{\Delta r_3}{\Delta r_2} c_2\right]\left[b_4 - \frac{\Delta r_4}{\Delta r_2} b_2\right] - \left[b_3 - \frac{\Delta r_3}{\Delta r_2} b_2\right]\left[c_4 - \frac{\Delta r_4}{\Delta r_2} c_2\right]}{\left[d_3 - \frac{\Delta r_3}{\Delta r_2} d_2\right]\left[b_4 - \frac{\Delta r_4}{\Delta r_2} b_2\right] - \left[b_3 - \frac{\Delta r_3}{\Delta r_2} b_2\right]\left[d_4 - \frac{\Delta r_4}{\Delta r_2} d_2\right]},$$

где

$$b_i = x_{ci} - x_{c1};$$

$$d_i = y_{ci} - y_{c1};$$

$$\rho_{ci} = x_{ci}^2 - y_{ci}^2;$$

$$c_i = \frac{1}{2}\left[\rho_{ci} - \rho_{c1} - \Delta r_i^2\right].$$

Используя спецификацию 3GPP TS 36.101 можно оценить, какую точность обеспечивает единичное значение параметра TA:

$$1TA = 16 \cdot T_{slot} = 16 \cdot \left(\frac{0.5\ \text{мс}}{15360}\right) = 0.52\ \text{мкс},$$

где T_{slot} – время одного таймслота. Зная значение TA (0,52 мкс) можно определить радиус окружности между базовой станцией и абонентским устройством:

$$1TA = \frac{3 \cdot 10^8 \cdot 0.52 \cdot 10^{-6}}{2} \approx 78 \text{ м.}$$

RTT (Round Trip Time) – это время между отправкой запроса и получением ответа. При помощи данного метода рассчитывается расстояние от базовой станции (БС) до БПЛА с использованием модели распространения сигнала (рисунок 2). Расстояние – это радиус круга вокруг базовой станции. Поэтому, если это возможно, для более точной оценки положения БПЛА следует использовать в вычислениях данные измерений времени RTT от соседних БС. Тогда БПЛА будет находиться в точке пересечения трех окружностей соседних БС, расположенных в центрах этих окружностей. Погрешность оценки местоположения БПЛА определяется пределами погрешностей всех измерений.

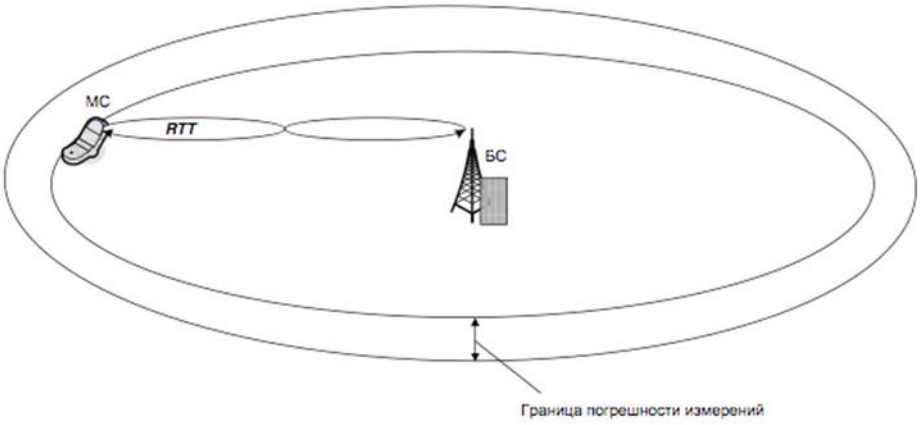


Рисунок 2 – Погрешность измерения методом RTT

Расстояние между БПЛА и БС можно выразить через время и скорость радиоволны определяется выражением:

$$D = \frac{RTT \cdot c}{2} + e,$$

где:

D – расстояние от БПЛА до БС;

c – скорость распространения радиоволны в соте, равная скорости света в свободном пространстве;

RTT – время обращения;

e – погрешность измерений.

В основе метода OTD (Observed Time Difference) – механизм синхронизации OTD-измерения и сравнения моментов прихода сигналов от разных БС на АТ. E-OTD (Enhanced OTD), специфицированный для GSM в 1999 году, является развитием механизма синхронизации OTD. Метод основан на измерении времени прохождения сигнала от БС до двух точек, первая из которых – БПЛА, а вторая – специальная станция измерения LMU (Location Measurement Unit), координаты которой известны. Различают гиперболический (разностно-дальномерный) TDoA/Hyperbolic и круговой (дальномерный) ToA/TDoA/Circular методы позиционирования E-OTD. Этот метод вводит следующие понятия временных разностей:

- OTD – наблюдаемая разность времени, которая замеряется БПЛА от двух БС в разных сотах. OTD = $t_2 - t_1$ (0 – если сигналы приходят одновременно)
- RTD – неточность временной синхронизации между двумя БС. RTD = $t_4 - t_3$ если БС передают сигнал в разные моменты времени и 0, если они синхронизированы
- GTD – геометрическая разность времени, которая обуславливается разным расстоянием от БПЛА до каждой БС. GTD = $(d_2 - d_1)/c$, при этом GTD связан с OTD и RTD: GTD = OTD – RTD

Метод Hyperbolic E-OTD проиллюстрирован на рисунке 3. OTD измеряется БПЛА, а RTD − LMU1, следовательно, GTD может быть вычислен. GTD21 определяет гиперболу $\Delta d_{21} = d_2 - d_1$, а GTD31 – гиперболу $\Delta d_{31} = d_3 - d_1$; пересечение GTD21 и GTD31 указывает на местоположение БПЛА. Для функционирования E-OTD необходимо наличие как минимум одного модуля LMU с известным местоположением для измерения параметра RTD рассинхронизации разных БС. Параметр GTD, характеризующий время распространения сигнала от БС к БПЛА, получается путем вычитания RTD из OTD. При оценке координат самим устройством БПЛА вычисляет свои координаты с использованием принимаемых из сети параметра RTD и известных координат БС. При оценке координат сетью БПЛА отправляет OTD, на основе которого координаты определяются в SMLC с использованием RTD и известных координат БС.

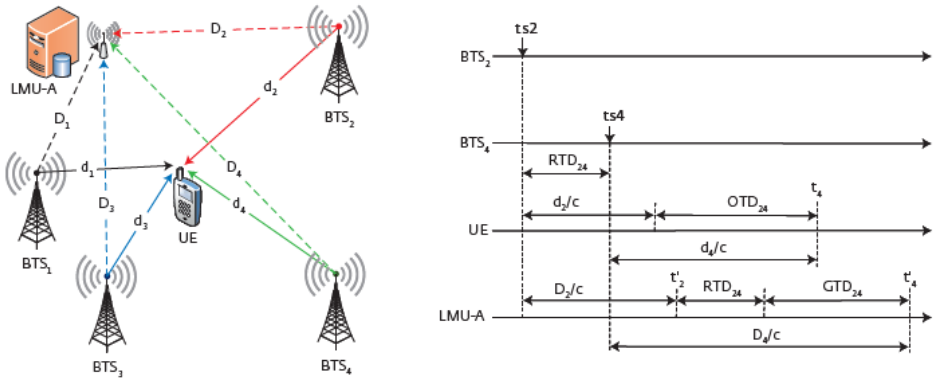


Рисунок 3 – Метод позиционирования E-OTD

Проиллюстрированный на рисунке 4 метод Circular E-OTD основан на измерении времени прихода сигнала в DL от БС до БПЛА и модуля измерения LMU, координаты которого точно известны. Для ОМП на плоскости необходимы как минимум три БС; часы БС, как и часы БПЛА и модуля измерения LMU, несинхронизированы.

Допустим, БСi излучает сигнал в канале BCCH/SCH. На БПЛА и модуле измерений LMU регистрируется время прихода сигнала от БСi: t_i – время прихода сигнала от БСi на БПЛА, T_i – время прихода сигнала от БСi на LMU. Пусть ε – рассинхронизация БПЛА и LMU, d_i – расстояние от БСi до БПЛА, а D_i – известное расстояние от БСi до LMU, тогда справедливо:

$$d_i - D_i = c(t_i - T_i + \varepsilon).$$

При оценке координат сетью значение t_i, зарегистрированное АТ, отправляется на модуль измерения LMU. Для позиционирования на плоскости эта операция повторяется минимум три раза с участием разных БС, после чего получается система уравнений:

$$d_1 - D_1 = c(t_1 - T_1 + \varepsilon),$$
$$d_2 - D_2 = c(t_2 - T_2 + \varepsilon),$$
$$d_3 - D_3 = c(t_3 - T_3 + \varepsilon),$$

Расстояния d_i можно выразить через координаты БПЛА (x, y), поэтому для ОМП БПЛА необходимо решить уравнение с тремя неизвестными: координаты БПЛА (x, y) и рассинхронизация ε.

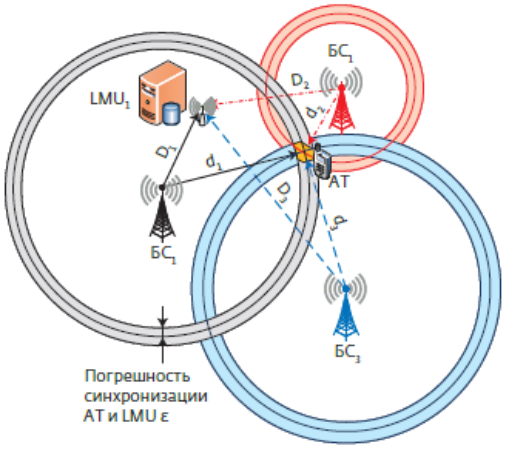


Рисунок 4 - Метод позиционирования Circular E-OTD

В методе AoA (Angels of Arrival) позиция БПЛА определяется пересечением двух линий, представляющих траектории контрольных сигналов, каждая из которых образует определенный угол с направлением от БПЛА на БС. Один измеренный угол образуется двумя линиями и дает возможность определить положение искомого беспилотника. Поэтому, как показано на рисунке 4, если БПЛА находится на линиях прямой видимости между двумя БС и проведено измерение углов AoA от двух соседних БС, то положение БПЛА находят на пересечении линий, определяемых углами AoA. Таким образом, можно вычислять местоположение БПЛА на основе информации о пересечении двух линий направления приема сигнала, т.е. эта технология требует, как минимум, двух БС, оборудованных антеннами с узкой диаграммой направленности. Большее количество используемых для расчетов станций даст лучшие показатели точности.

Эта технология находит применение в сетях, построенных с использованием секторных и адаптивных антенн (антенных решеток) с узкой диаграммой направленности (ДН). Точность технологии AoA на практике не высока, кроме того, существуют проблемы практического характера. Во-первых, точность определения направления приема прямо пропорциональна ширине ДН применяемой антенны. Во-вторых, из-за эффекта многолучевого распространения радиоволн БС не всегда может отличить прямой сигнал БПЛА от переотраженных. Эта проблема не является специфической для технологии AoA и характерна для большинства технологий определения местоположения. Точность данного метода может быть повышена с увеличением числа измерений, что требует дополнительной производительности сети. К недостаткам метода можно отнести его сложность и дороговизна, так как стоимость секторных и адаптивных антенн значительно превышает стоимость обычных антенн.

Сравнительный анализ технологий определения местоположения в сетях LTE может быть проведен по нескольким параметрам:

- точность определения местоположения;
- стоимость внедрения;
- нагрузка на сеть – дополнительная нагрузка на отдельные подсистемы сети (HLR, коммутаторы и т.п.);
- возможный режим работы платформы:
- пассивный – платформа накапливает в себе информацию о местоположении абонентов и при запросе выдаёт её;

• активный – при запросе на определение местоположения платформа извлекает информацию о местоположении абонента или данные, необходимые для его расчета;

• поддержка со стороны абонентского устройства – необходимо ли наличие в мобильном терминале и/или СИМ-карте аппаратного/программного блока, необходимого для работы системы определения местоположения.

Результаты анализа вышеуказанных методов по данным критериям представлены в таблице 1.

Таблица 1 – Результаты анализа методов обнаружения

	Cell ID	**Cell ID+TA**	**E-OTD**	**RTT**
Точность определения местоположения	20-150 м	6-35 м	высокая	высокая
Стоимость внедрения	низкая	низкая	средняя	средняя
Нагрузка на сеть	низкая	низкая	средняя	высокая
Режим работы платформы - **а**ктивный, **п**ассивный	а,п	а,п	а,п	а,п
Требует поддержки со стороны телефона	нет	нет	да	да

Литература

1. Chirov D.S., Chertova O. G.; Rymshin K. Y. Covert Detection of Air Targets with a Passive Radar System Using External Signals from Broadcast TV Stations of the DVB-T2 Standard // 25th International Conference on Digital Signal Processing and its Applications (DSPA), Moscow, Russian Federation, 2023, pp. 1-5, doi: 10.1109/DSPA57594.2023.10113448.
2. Рюмшин К.Ю., Кочетков Ю.А., Атакищев О.И., Чиров Д.С. Алгоритм и результаты моделирования процесса обработки сигналов в многопозиционной радиолокационной станции обнаружения беспилотных летательных аппаратов // Известия ИИФ. № 2(64), 2022. С.24-34.
3. Киреев А. В., Фокин Г. А. Пеленгация источников радиоизлучения LTE мобильным пунктом радиоконтроля с круговой антенной решеткой // Труды Научно-исследовательского института радио. 2015. № 2. С. 68–71.
4. Sivers M., Fokin G. LTE Positioning Accuracy Performance Evaluation // Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer International Pub-lishing, 2015. С. 393–406.
5. Сиверс М. А., Фокин Г. А., Духовницкий О. Г. Позиционирование абонентских станций в сетях мобильной связи разностно-дальномерным методом // Системы управления и информационные технологии. 2015. Т. 59. № 1. С. 55–61.

ANALISE OF POSSIBILITES OF DETERMINATION COORDINATES UAV BY ONBOARD LTE MODEM SIGNALS

stud. Ryabchitsky M.M., stud. Illarionova P.V.

Moscow technical university of communication and informatics

A huge need in tracking different unmanned vehicles and their movement trajectories take place in todays reality, along with UAV identification and classification among others network subscribers. This article reviews different methods of determining coordinates of UAV by onboard LTE modem signal, through base stations. Each method is described in details, along with it’s realization and advantages and disadvantages, such as: realization difficulty, accuracy, reliability and cost.

КЛАСТЕРНЫЙ АНАЛИЗ ПУБЛИКАЦИЙ ПО БЕСПРОВОДНОЙ СВЯЗИ НА ОСНОВЕ СИГНАЛОВ С ОРБИТАЛЬНО-МОДОВОЙ МОДУЛЯЦИЕЙ

асп. Орешкина А.К., к.т.н. Петров С.А.

ФГУП Всероссийский научный исследовательский институту автоматики им. Духова

В работе выполнен кластерный анализ научных публикаций, посвящённых применению сигналов с орбитально-модовой модуляцией в системах беспроводной связи, с целью оценки структуры и степени сформированности исследований для подводных каналов передачи данных. Сформирован двуязычный корпус и построена сеть тематического сходства на основе семантических эмбеддингов текстов. Кластеризация методом оптимизации модульности позволила выделить ключевые направления и их взаимосвязи. Показано, что исследования преимущественно сосредоточены в оптическом и радиочастотном диапазонах, тогда как гидроакустические решения представлены ограниченно. Результаты кластерного анализа позволили систематизировать существующие технические решения и определить их применимость в качестве основы для проектирования гидроакустических систем связи с использованием сигналов с орбитально-модовой модуляцией.

Введение

Современные системы беспроводной связи, в том числе применяемые в подводных робототехнических комплексах (РТК), характеризуются возрастающими требованиями к пропускной способности и скорости передачи данных, особенно при работе на малых и средних дистанциях. Неоднородность гидроакустической (ГА) среды, ограниченная полоса частот и выраженные искажения канала приводят к снижению эффективности традиционных методов модуляции и сдерживают рост производительности таких систем [1]. В этих условиях перспективным направлением является использование сигналов с орбитально-модовой модуляцией, основанных на применении орбитального углового момента волны (ОУМ) и обеспечивающих пространственное мультиплексирование за счёт ортогональных пространственных мод. Реализация дополнительной пространственной степени свободы потенциально позволяет повысить спектральную эффективность канала, что особенно актуально для РТК, требующих передачи видеоинформации, телеметрии и управляющих команд в реальном времени.

Концепция использования ОУМ волн в задачах связи активно развивается с начала 2010-х годов, прежде всего в оптических и радиочастотных системах, где продемонстрированы возможности пространственного мультиплексирования независимых мод и увеличения пропускной способности каналов передачи данных [2]. Перенос данных подходов в ГА среду требует учёта специфики акустического канала распространения, включая конструктивные ограничения излучающих и приёмных массивов, выраженную многолучевость, фазовые искажения и доплеровские эффекты, способные нарушать ортогональность пространственных мод. В связи с этим для оценки текущего состояния исследований и определения перспектив применения орбитально-модовой модуляции в гидроакустике представляется целесообразной систематизация накопленных научных результатов.

Целью работы является выявление структуры и степени сформированности исследований, посвящённых применению орбитально-модовой модуляции в ГА канале, на основе кластерного анализа научных публикаций. Для достижения поставленной цели формируется совокупность релевантных источников, выполняется их тематическая кластеризация, определяются ключевые направления и взаимосвязи между ними. Полученные результаты позволяют оценить текущее состояние области, выделить перспективные направления развития и обосновать дальнейшие подходы к построению высокопроизводительных подводных систем связи.

Основные принципы построения гидроакустических систем связи, и оценка возможности применения ОУМ

Типовая ГА система связи включает передающий и приёмный тракты, соединённые ГА каналом распространения, и в общем виде представлена на рисунке 1. В общем виде передающий тракт содержит: источник информации (ИИ), кодер, модулятор, полосовой фильтр (ПФ), усилитель мощности (УМ) и антенну.

В приёмном тракте последовательно применяются УМ, ПФ, демодулятор и декодер.

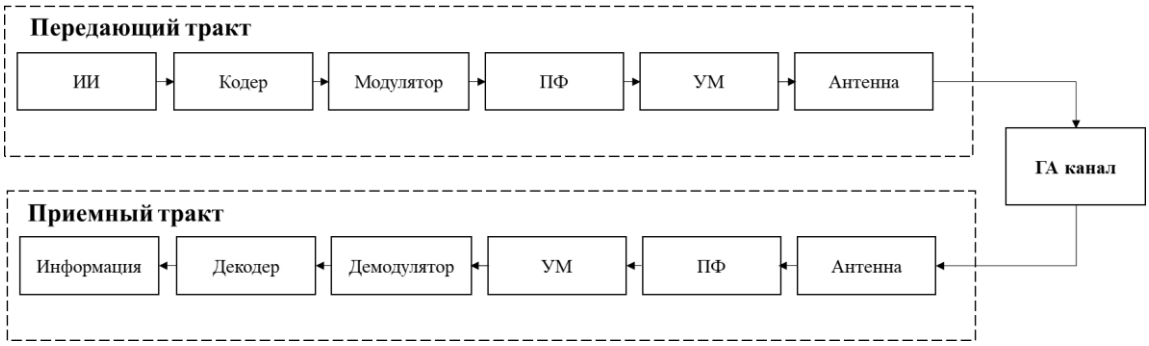


Рисунок 1– Типовая система ГА связи

Передача осуществляется преимущественно в низкочастотном диапазоне с использованием частотной, фазовой или многочастотной модуляции, при этом фактически задействуется один пространственный канал распространения [3]. Реальный ГА канал характеризуется многолучевостью вследствие отражений от поверхности и дна, фазовыми искажениями, обусловленными неоднородностью профиля скорости звука, а также доплеровскими смещениями [1]. Указанные эффекты приводят межсимвольной интерференции и снижению отношения сигнал/шум. Ограниченная полоса частот и значительные потери распространения обуславливают сравнительно низкие скорости передачи данных, как правило, единицы–десятки кбит/с [4].

Одним из направлений повышения пропускной способности является пространственное мультиплексирование, в частности применение сигналов с орбитально-модовой модуляцией, основанных на использовании ОУМ волны. Акустическая вихревая мода описывается выражением:

$$u_l(r) \propto e^{il\varphi},$$

где φ — азимутальная координата, l — топологический заряд.

Ортогональность мод с различными значениями l позволяет рассматривать их как независимые пространственные каналы. Передаваемое поле при использовании нескольких мод представляется суперпозицией:

$$p(r,t) = \sum_{l \in L} s_l(t) u_l(r),$$

где $s_l(t)$— сигналы отдельных потоков.

В идеальном случае пропускная способность системы возрастает пропорционально числу задействованных мод.

В реальном ГА канале неоднородность среды и многолучевое распространение искажают пространственную структуру поля и могут вызывать смешение мод, что усложняет их формирование и разделение на приёмной стороне. В литературе предлагаются различные подходы к реализации и обработке сигналов с орбитально-модовой модуляцией, однако данные решения существенно различаются по применяемым методам и техническим средствам. Это делает целесообразной систематизацию существующих исследований и анализ их структуры.

Методика кластерного анализа

Для выявления тематических кластеров научных публикаций применён сетевой анализ с использованием программного обеспечения Gephi и алгоритма оптимизации модулярности Louvain [5].

Первый этап включал сбор и подготовку данных, в ходе которого был сформирован двуязычный корпус публикаций. Поиск осуществлялся в базах Web of Science, Scopus и eLIBRARY.ru по обобщённым тематическим запросам, включающим сочетания терминов, связанных с орбитально-модовой модуляцией и задачами беспроводной связи (например: «orbital angular momentum» AND «communication»; русскоязычный вариант: «орбитальный угловой момент» И «связь»). Такой подход обеспечивал максимально широкий охват исследований без предварительного ограничения по типу канала распространения или физической природе среды. Это позволило сформировать репрезентативный корпус работ и выполнить последующее выделение специализированных направлений, включая ГА применения, уже на этапе кластерного анализа.

Таким образом англоязычный корпус включал 224 публикаций за период 2010–2025 гг., что соответствует требованиям репрезентативности для сетевых и библиометрических методов анализа [6]. Русскоязычный корпус формировался по аналогичным критериям и составил 111 публикаций.

На втором этапе для каждой публикации формировался текстовый дескриптор, объединяющий заголовок, аннотацию и авторские ключевые слова. Тексты очищались от служебных символов и стоп-слов и приводились к нормализованному виду. При отсутствии авторских ключевых слов они автоматически извлекались методом KeyBERT [7], основанным на семантических эмбеддингах BERT.

Для представления дескрипторов в числовом пространстве применялись контекстные трансформерные модели Sentence-BERT [8]. Для англоязычных текстов использовалась специализированная модель SPECTER [9], обученная на корпусе научных статей с учётом структуры цитирования. Для русскоязычных текстов применялась мультиязычная модель *paraphrase-multilingual-mpnet-base-v2*. Каждому дескриптору сопоставлялся нормированный вектор признаков, а семантическая близость между публикациями вычислялась по косинусной мере сходства.

Третий этап заключался в построении взвешенной сети на основе матрицы попарных значений сходства: узлы сети соответствуют публикациям, а рёбра формируются между теми из них, чьё значение сходства превышает заданный порог. Пороговые значения выбирались по анализу распределения коэффициентов сходства, отражающего статистику семантической близости статей внутри корпуса. Полученные сообщества интерпретировались как устойчивые тематические направления исследований.

Ограничения и допущения

Предложенная методика имеет ряд ограничений, обусловленных как особенностями исходных данных, так и используемыми инструментами обработки.

Во-первых, значительная часть публикаций в наукометрических базах недоступна в полнотекстовом виде, что обусловило использование только открытых текстовых дескрипторов. Такой подход является общепринятым в библиометрических исследованиях [6], однако может приводить к частичной потере семантической информации и снижению точности тематического сопоставления.

Во-вторых, результаты анализа зависят от выбранных параметров векторизации и кластеризации, включая используемые языковые модели, косинусную меру сходства и порог формирования рёбер графа. Изменение данных параметров способно влиять на плотность сети и границы выделяемых сообществ, что может приводить к вариациям структуры кластеров.

В-третьих, тематическая кластеризация отражает прежде всего семантическую близость текстовых описаний публикаций и не учитывает напрямую экспериментальные результаты, техническую реализуемость или уровень зрелости конкретных решений. Следовательно, полученные кластеры интерпретируются как индикаторы направлений исследований, а не как количественная оценка эффективности технологий.

Результаты

Семантическая сеть, построенная для англоязычного корпуса, характеризуется выраженной модульной структурой: значение модульности составляет 0,34, что свидетельствует о наличии отчётливо разделённых тематических сообществ.

Алгоритмом оптимизации модулярности выявлено 5 кластеров. Их содержательная интерпретация приведена в таблице 1.

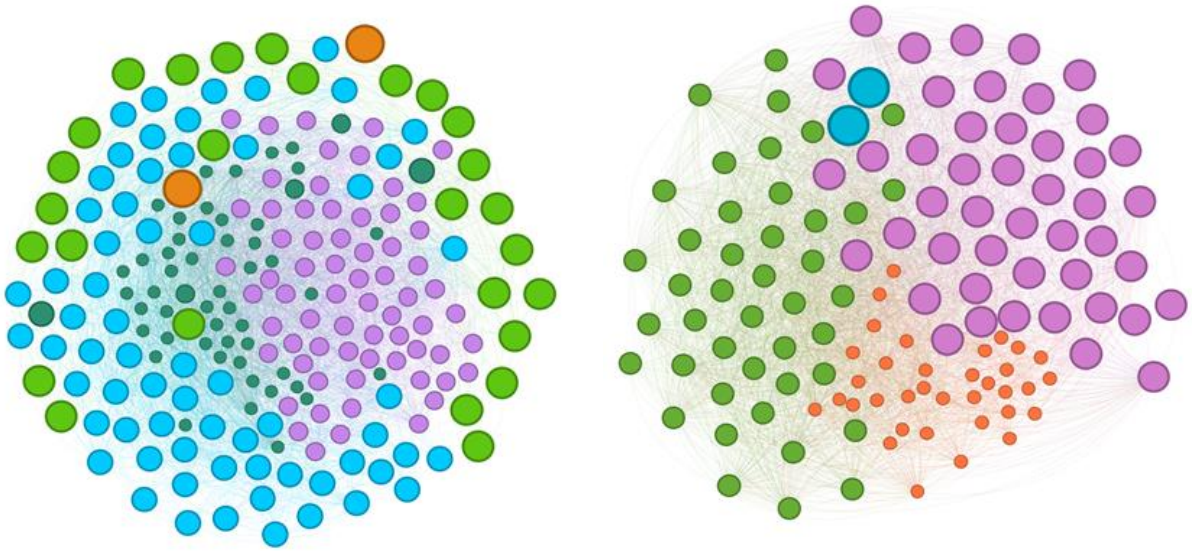
Согласно визуализации сети (рисунок 2а), подводный кластер формирует компактную группу на периферии графа. При этом он сохраняет связи с кластерами генерации ОУМ-мод и беспроводных систем, что указывает на использование общих технических решений и методов распространения волн. Такое положение подтверждает тематическую самостоятельность данного направления при одновременной междисциплинарной связанности с базовыми исследованиями в области ОУМ коммуникаций

Для русскоязычного корпуса семантическая сеть публикаций построена по аналогичной методике. Полученная структура характеризуется менее выраженной сегментацией. Значение

модулярности составило 0,15, что указывает на слабую разделённость сети и более высокую тематическую однородность корпуса.

Алгоритмом выявлено 4 кластера (рисунок 2б). Их содержательная интерпретация представлена в таблице 2. Тематически кластеры преимущественно связаны с задачами фундаментальной оптики вихревых и Лагерра-Гауссовых пучков, моделированием их распространения и исследованием характеристик ОУМ.

В отличие от англоязычного корпуса, исследования, ориентированные непосредственно на задачи передачи информации и связи, не формируют самостоятельного крупного сообщества и представлены небольшими слабо изолированными подгруппами. Основная часть публикаций относится к теоретическим и экспериментальным работам по фундаментальной оптике.



а) б)

Рисунок 2– Семантическая сеть публикаций а) англоязычного корпуса, б) русскоязычного корпуса

Таблица 1. Тематические кластеры семантической сети публикаций англоязычного корпуса

Цвет кластера	**Тема**	**Количество публикаций**
Фиолетовый	Теоретические и алгоритмические методы обработки ОУМ-сигналов	70
Голубой	Пространственное мультиплексирование и беспроводные ОУМ-системы	68
Зеленый	Формирование и генерация ОУМ-мод	56
Салатовый	Подводная связь	28
Оранжевый	Малый экспериментальный физический кластер	2

Таблица 2. Тематические кластеры семантической сети публикаций русскоязычного корпуса

Цвет кластера	**Тема**	**Количество публикаций**
Фиолетовый	Физика и свойства вихревых пучков	47
Зеленый	Формирование и моделирование ОУМ-пучков	41
Оранжевый	Теоретические основы	21
Синий	Распространение ОУМ-волн в турбулентной среде	2

Анализ подводного кластера

Анализ публикаций, относящихся к подводному направлению передачи данных с использованием сигналов с орбитально-модовой модуляцией, показал, что основная часть исследований в акустической части касается экспериментальной генерации акустических вихревых мод и демонстрации их пригодности для передачи информации. Показано, что акустические вихревые лучи с разными топологическими зарядами могут быть сгенерированы фазированной решёткой и использованы для передачи ASCII-кодированной информации, что подтверждает фундаментальную применимость ОУМ в гидроакустике для увеличения числа независимых каналов связи на одной несущей частоте благодаря ортогональности мод. Более новые работы экспериментально исследуют схемы мультиплексирования и демультиплексирования таких мод, в том числе с использованием кольцевых массивов излучателей и гидрофонов, а также алгоритмов пространственной декомпозиции поля для точного извлечения информации из комбинации ОУМ-компонент даже в частично измеренных полях, что демонстрируется в современных исследованиях, предлагающих динамическую модальную декомпозицию как метод эффективной детекции топологического заряда и снижения ошибок.

В этих работах формирование акустических вихревых мод осуществляется через фазированные излучающие решётки, где цифровое управление фазами каждого излучателя задаёт амплитудно-фазовую картину выходного поля, реализуя пространственный ОУМ-модулятор. На передающей стороне источником информации является цифровой поток, кодируемый в набор фазовых и амплитудных значений, распределённых по модам. Усиление мощности и фильтрация осуществляются традиционными методами усиления сигнала и полосовой фильтрации в окрестности несущей частоты, а сам многоканальный излучающий массив служит одновременно модулятором и антенной.

На приёмной стороне гидрофонный массив служит для пространственной выборки акустического поля, после чего применяются полосовые фильтры и усилители для подготовки сигналов к цифровой обработке. Основной демодулирующий блок представляет собой пространственную демультиплексацию мод, реализуемую с помощью методов модальной декомпозиции или алгоритмов динамической модальной декомпозиции, что позволяет выделить отдельные ОУМ-каналы и затем выполнить традиционную цифровую демодуляцию и декодирование данных.

Параллельно исследования в области подводной оптической связи с использованием ОУМ показывают более зрелую реализацию концепции пространственного мультиплексирования, где лазерный источник и электро-оптические модуляторы формируют несколько независимых ОУМ-лучей, каждый из которых несёт собственный поток данных, а приёмная оптика и DSP-алгоритмы выполняют демультиплексирование и компенсацию межмодовых искажений. Эти оптические архитектуры демонстрируют, что ОУМ может быть интегрирован с существующими схемами модуляции (OFDM/QAM), а методы цифровой обработки из оптики потенциально могут быть адаптированы для ГА систем.

Заключение

В работе предложена методика кластерного анализа научных публикаций, посвящённых применению сигналов с орбитально-модовой модуляцией в системах связи. На основе двуязычного корпуса и графа тематического сходства, сформированного с использованием семантических эмбеддингов, кластеризация методом оптимизации модульности позволила выделить устойчивые направления исследований и получить системный, воспроизводимый обзор предметной области.

Показано, что большинство работ сосредоточено в оптическом и радиочастотном диапазонах, где ОУМ используется как дополнительная пространственная степень свободы для мультиплексирования каналов. ГА решения представлены ограниченно и носят преимущественно экспериментальный характер, однако подтверждают применимость аналогичных принципов: формирование мод фазированными излучающими решётками, приём гидрофонными массивами и последующую цифровую пространственную демультиплексацию с использованием ортогонального декодирования и алгоритмов динамической модальной декомпозиции.

Полученные результаты формируют основу для дальнейшей разработки ГА систем, использующих пространственную степень свободы для повышения пропускной способности.

Литература

1. Урик Э. Л. Основы гидроакустики / Э. Л. Урик. – 2-е изд., перераб. и доп. – Л. : Судостроение, 1978. – 448 с
2. Bozinovic N. Terabit-scale orbital angular momentum mode division multiplexing in fibers / N. Bozinovic, Y. Yue, Y. Ren, M. Tur, P. Kristensen, H. Huang, A. E. Willner, S. Ramachandran // Science. – 2013. – Vol. 340, № 6140. – P. 1545–1548. – DOI: 10.1126/science.1237861.
3. Stojanovic M. Underwater acoustic communication channels: propagation models and statistical characterization / M. Stojanovic, J. Preisig // IEEE Communications Magazine. – 2009. – Vol. 47, № 1. – P. 84–89. – DOI: 10.1109/MCOM.2009.4752682.
4. Kilfoyle D. The state of the art in underwater acoustic telemetry / D. Kilfoyle, A. Baggeroer // IEEE Journal of Oceanic Engineering. – 2000. – Vol. 25, № 1. – P. 4–27. – DOI: 10.1109/48.820733.
5. Blondel V. D. Fast unfolding of communities in large networks / V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre // Journal of Statistical Mechanics: Theory and Experiment. – 2008. – № 10. – P. P10008. – DOI: 10.1088/1742-5468/2008/10/P10008.
6. Donthu N. How to conduct a bibliometric analysis: an overview and guidelines / N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, W. Lim // Journal of Business Research. – 2021. – Vol. 133. – P. 285–296. – DOI: 0.1016/j.jbusres.2021.04.070.
7. Grootendorst M. KeyBERT: minimal keyword extraction with BERT [Электронный ресурс] / M. Grootendorst. – GitHub repository, 2020. URL: https://github.com/MaartenGr/KeyBERT (дата обращения: 01.02.2026).
8. Reimers N. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks / N. Reimers, I. Gurevych // Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. – 2019. – P. 3982–3992. – DOI: 10.18653/v1/D19-1410.
9. Cohan A. SPECTER: Document-level representation learning using citation-informed transformers / A. Cohan, S. Feldman, I. Beltagy, D. Downey, D. S. Weld // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. – Association for Computational Linguistics, 2020. – P. 2270–2282. – DOI: 10.18653/v1/2020.acl-main.207.

CLUSTER ANALYSIS OF PUBLICATIONS ON WIRELESS COMMUNICATION BASED ON ORBITAL-MODE MODULATION SIGNALS

Oreshkina A.K., Petrov S.A.

FSUE VNIIA

The paper performs a cluster analysis of scientific publications on the application of signals with orbital-mode modulation in wireless communication systems, in order to assess the structure and degree of formation of research for underwater data transmission channels. A bilingual corpus is formed and a network of thematic similarity is built based on semantic embeddings of texts. Clustering by the method of modularity optimization allowed to identify key directions and their interrelations. It is shown that research is predominantly concentrated in the optical and radio frequency bands, while hydroacoustic solutions are presented in a limited way. The results of the cluster analysis made it possible to systematize the existing technical solutions and determine their applicability as a basis for designing hydroacoustic communication systems using signals with orbital-mode modulation.

РАЗРАБОТКА АРХИТЕКТУРНОГО ПОДХОДА К ОБНАРУЖЕНИЮ СЕТЕВЫХ АНОМАЛИЙ В РАЗНОРОДНЫХ МАСШТАБНЫХ СЕТЯХ IOT

к.т.н., доц. Сорокин И.А., преп. Чесноков А.Д.

Нижегородский государственный инженерно-экономический университет

Статья посвящена разработке архитектуры кластеризованного обучения для обнаружения сетевых аномалий в крупномасштабных гетерогенных сетях Интернета вещей (IoT). Актуальность исследования обусловлена экспоненциальным ростом векторов атак и низкой эффективностью традиционных сигнатурных методов защиты в условиях полиморфизма вредоносного ПО и высокой гетерогенности устройств. Авторами предлагается подход, основанный на кластеризованном обучении, который позволяет обучать неконтролируемые модели (автоэнкодеры) на распределённых данных без их централизованного сбора, обеспечивая конфиденциальность и снижая сетевую нагрузку. Ключевым элементом архитектуры является интеграция неконтролируемого алгоритма кластеризации устройств, анализирующего параметры частично обученных моделей для формирования адаптивных глобальных моделей, специфичных для групп устройств со схожими паттернами поведения.

Введение

Широкомасштабное проникновение технологий Интернета вещей в корпоративный и промышленный секторы обеспечивает качественный скачок операционной эффективности за счёт автоматизации процессов, снижения издержек и минимизации влияния человеческого фактора. Подобные улучшения становятся возможными благодаря конвергенции киберфизических систем и сетевых инфраструктур, позволяющих устройствам автономно собирать и обрабатывать данные для управления физическими процессами [1].

Однако глубокая интеграция и гетерогенность IoT-устройств неизбежно ведут к экспоненциальному росту векторов атак и повышению уязвимости как потребительского, так и индустриального сегментов. Ситуация усугубляется рыночным трендом на использование готовых коммерческих компонентов, где приоритет отдаётся скорости вывода продукта и простоте эксплуатации, а не защищённости.

Совокупность программных уязвимостей [5], недостаточная эффективность существующих средств защиты и повсеместная связанность устройств формируют благоприятную среду для киберпреступников. Классическим примером является ботнет Mirai, который эксплуатировал проблему стандартизированных и слабых учётных данных для массовой компрометации разнородных устройств. Современные угрозы демонстрируют эскалацию сложности: модификации Mirai, появившиеся после утечки исходного кода, наряду с продвинутым вредоносным ПО, используют комплексные методы эксплуатации сетевых протоколов и программных ошибок. Скомпрометированные узлы традиционно применяются для организации DDoS-атак, скрытого майнинга, спам-рассылок и фрод-активности.

Особую обеспокоенность вызывает безопасность промышленного Интернета вещей. Атаки на критические объекты инфраструктуры несут риски техногенных катастроф, саботажа производства, хищения интеллектуальной собственности и использования заражённых узлов как плацдарма для проникновения в корпоративные сети и технологические сети.

Обсуждение

Традиционные подходы к защите, включающие ужесточение конфигураций и управление обновлениями, не гарантируют полной безопасности из-за человеческих ошибок, появления уязвимостей нулевого дня и сложности конфигурирования [2]. Системы обнаружения и предотвращения вторжений, основанные на сигнатурном анализе, демонстрируют снижение эффективности перед лицом полиморфизма вредоносного кода, обфускации и динамической инфраструктуры ботнетов. Задержки в актуализации сигнатурных баз делают такие системы неоперативными против новых угроз [3].

Применение методов машинного и глубокого обучения открывает новые горизонты в обнаружении аномалий благодаря их адаптивности и обобщающей способности. Тем не менее, централизованная парадигма обучения ML-моделей сталкивается с критическими ограничениями в масштабных и гетерогенных сетях IoT. К ним относятся: чрезмерная нагрузка на каналы связи, высокие задержки и потери пакетов при передаче телеметрии в облако. Помимо технических проблем, централизация данных порождает вопросы соблюдения

конфиденциальности и соответствия регуляторным нормам. Периферийные вычисления частично снимают нагрузку с магистральных сетей, но порождают проблему фрагментации данных, ограничивая репрезентативность обучающих выборок [4].

В качестве парадигмального решения, одновременно адресующего проблемы сетевой нагрузки, приватности и изоляции данных, выступает кластеризированное обучение. Оно позволяет обучать общую модель на распределённых данных без их централизованного сбора. Клиенты выполняют локальное обучение и передают на сервер только обновления весов, что гарантирует конфиденциальность первичных данных и кратно сокращает сетевой трафик.

Однако практическая имплементация в гетерогенных средах сопряжена с проблемой статистической гетерогенности: распределение данных на различных клиентах может существенно отличаться от независимого и одинаково распределённого, что препятствует сходимости единой глобальной модели. Данная проблема наиболее остро проявляется в крупномасштабных сетях IoT, характеризующихся множеством протоколов и паттернов поведения.

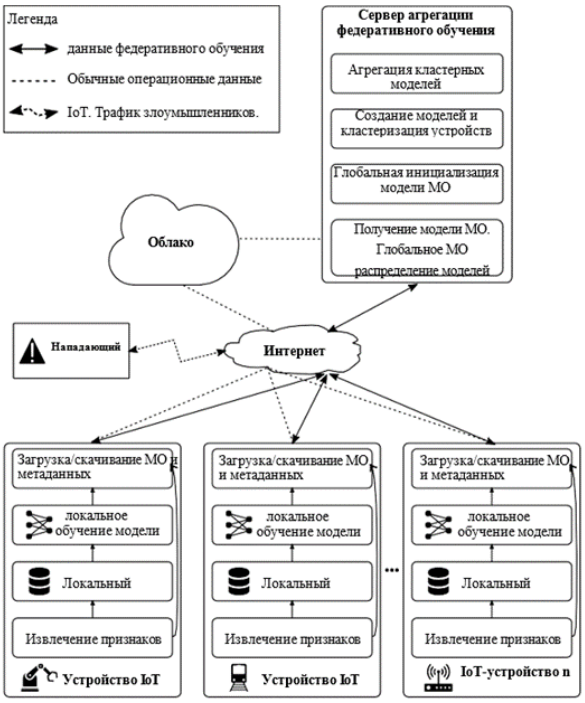


Рисунок 1. Предлагаемая архитектура системы.

Предлагаемая архитектура для обучения IDS изображена на рисунке 1. Она включает множество клиентов и центральный сервер агрегации, а также показывает потенциальных атакующих.

Предлагаемая система предназначена для работы в большой сети гетерогенных устройств IoT, таких как шлюзы, CPS и промышленные машины, которые общаются с использованием разных протоколов. Эти устройства могут располагаться в разных сегментах сети или быть географически распределены, что может влиять на их поведение. Устройства постоянно отправляют/получают данные в/из облачного слоя.

Каждое устройство считается клиентом в процессе раздельного обучения. Они отвечают за захват релевантных данных, локальное обучение модели ML и вывод модели для обнаружения аномалий после завершения обучения. Никакие обучающие данные не передаются на сервер агрегации, только параметры модели и минимальные метаданные, релевантные для процесса раздельного обучения. Ожидается, что устройства будут выполнять легковесные задачи ML, но для низкомощных устройств IoT, таких как датчики и актуаторы, система ожидается развернутой на уровне хаба или шлюза. В этой работе отдается приоритет использованию легковесных

моделей ML для обнаружения аномалий, чтобы ограничить вычислительную нагрузку во время обучения модели или вывода.

Сервер агрегации координирует весь процесс обучения, инициализируя и распределяя модель и гиперпараметры обучения клиентам, получая обновления модели от клиентов, выполняя дактилоскопию модели и кластеризацию устройств, запуская агрегацию по кластерам полученных моделей и отправляя соответствующую агрегированную глобальную модель каждому клиенту.

В этой статье рассматривается две основные модели угроз. Первая учитывает внешних актеров, которые удаленно сканируют устройства IoT в сети, находят уязвимые устройства для эксплуатации и удаленно компрометируют их. Вторая предполагает локального противника, компрометирующего одно или несколько устройств IoT в защищенной сети и использующего их для запуска атак против других устройств в той же сети или нацеливания на внешних жертв.

Для развертывания рассматривается типичная настройка кросс-девайс с большим количеством устройств. Однако из-за гарантий доступности, требуемых многими устройствами IoT, особенно в промышленных настройках, ожидается, что большинство устройств будут участвовать в процессе раздельного обучения. Это позволяет серверу поддерживать постоянное состояние для каждого клиента для выполнения шага кластеризации. В этой работе мы предполагаем, что ни одно устройство IoT не заражено до обучения модели и что ни одно из них не ведет себя враждебно.

Предлагаемое кластеризованное обучение работает следующим образом. Сначала сервер агрегации инициализирует веса модели W_0 и выбирает гиперпараметры обучения. Затем сервер отправляет эти значения всем участвующим клиентам. На следующем шаге каждый клиент частично обучает W_0, используя только свои локальные данные, в течение определенного количества эпох. Локальное обучение выполняется с использованием оптимизатора на основе градиента CLIENTOPT для минимизации локальных потерь обучения. CLIENTOPT является абстракцией для оптимизаторов, таких как SGD, Adam или RMSprop.

Для каждого идентифицированного кластера k выполняется независимый процесс параллельно. В ходе работы осуществляется несколько раундов (всего R раундов) до тех пор, пока глобальная модель для каждого кластера не сойдётся – в итоге формируется набор из K глобальных моделей. В рамках каждого раунда клиенты передают на сервер разницу между весами полученной в начале раунда глобальной модели и весами, обновлёнными локально. Сервер, в свою очередь, использует эти дельты весов для вычисления так называемого псевдоградиента – отрицательного среднего значения дельт моделей. Псевдоградиент совместно со скоростью обучения сервера ηs задействуется в процессе агрегации модели, который обобщён в функции SERVEROPT.

Абстракция CLIENTOPT и SERVEROPT даёт возможность интегрировать импульс или иные адаптивные методы оптимизации как на стороне клиента (в процессе обучения), так и на стороне сервера (в процессе агрегации модели) – в отличие от алгоритма FedAvg. При этом популярный метод агрегации FedAvg можно рассматривать как частный случай, при котором SERVEROPT настроен на SGD со скоростью обучения сервера ηs=1,0. В сети гетерогенных устройств базовое распределение данных зачастую не является независимым и одинаково распределенным.

В контексте единственная глобальная модель достаточной сложности теоретически могла бы адекватно подойти к данным, однако обучение столь сложной модели на устройствах IoT нередко оказывается невозможным из-за аппаратных ограничений. По этой причине предлагается группировать устройства со схожим поведением, чтобы создать набор глобальных моделей, специально адаптированных для каждой группы. В рамках данного подхода каждое устройство IoT получает метку группы в неконтролируемом режиме – эта метка будет использоваться в процессе раздельного обучения.

Использование локально обученных обновлений модели в качестве входов для метода кластеризации обладает несколькими ключевыми преимуществами: во-первых, отсутствует необходимость интегрировать внешние алгоритмы дактилоскопии устройств или применять ручные методы; во-вторых, не требуется выжидать определённое время для идентификации устройств перед стартом обучения модели; в-третьих, весь процесс полностью интегрирован в конвейер раздельного обучения.

Первый этап дактилоскопии модели заключается в частичном обучении каждой локальной модели в течение ϵ эпох с последующей отправкой частично обученной модели на сервер агрегации. Далее сервер выравнивает параметры (веса и смещения) каждой модели и выполняет анализ главных компонент (PCA) для снижения размерности параметров. Снижение размерности позволяет ускорить вычисления на этапе кластеризации, а также помогает снизить риски, связанные с кластеризацией данных высокой размерности в моделях, содержащих значительное число параметров. Для кластеризации данных сниженной размерности применяется алгоритм K-means с схемой инициализации k-means++.

Гипотеза заключается в том, что клиенты с похожими распределениями данных сойдутся к моделям со схожими значениями параметров (весов и смещений) при условии, что все клиенты начинают с одной и той же начальной случайной модели W_0. Поскольку предложение носит неконтролируемый характер, для выбора оптимального значения количества кластеров K будут использоваться внутренние метрики валидации кластеризации. Эти метрики не опираются на внешние данные и в основном базируются на таких показателях, как компактность образцов внутри одного кластера и разделение между разными кластерами. В частности, для выбора значения K будут оценены следующие внутренние метрики валидации: Silhouette, Davies-Bouldin и SDbw баллы.

Материалы и методы

В рамках данной статьи в качестве моделей обнаружения аномалий используются автоэнкодеры нейронных сетей, которые уже находили применение в схожих доменах для выявления сетевых атак. Приоритет отдаётся легковесным автоэнкодерам (с небольшим числом параметров), что особенно актуально для рассматриваемого сценария развёртывания. Такой выбор обусловлен несколькими причинами: во-первых, снижается вычислительная нагрузка при обучении модели и выполнении выводов на ограниченных устройствах; во-вторых, уменьшается объём сетевого трафика между устройствами и сервером агрегации в ходе раундов – за счёт меньшего числа параметров по сравнению с более сложными моделями.

Автоэнкодеры представляют собой неконтролируемые нейронные сети, стремящиеся воспроизвести входные данные на выходном слое при соблюдении определённых ограничений, что позволяет избежать обучения функции идентичности. Структура автоэнкодера включает две сети: энкодер и декодер. Энкодер принимает входные признаки $x \in \mathrm{R}n$ и преобразует их в скрытое закодированное пространство $h \in \mathrm{R}e$, где $e<n$ (это ограничение предотвращает обучение функции идентичности). Далее декодер преобразует h в $x' \in \mathrm{R}n$. Цель автоэнкодера заключается в минимизации среднеквадратичной ошибки (MSE, ошибки реконструкции) между x и x' – как указано в уравнении (1).

Автоэнкодер обучается с использованием функции потерь, показанной в уравнении (2), которая в дополнение к MSE включает член регуляризации L2.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - x_i'\right)^2. \quad (1)$$

$$\zeta = MSE + \lambda \sum_i \omega_i^2 \quad (2)$$

Обучается автоэнкодер с использованием образцов нормального (легитимного или benign) сетевого трафика IoT, который не содержит атак; таким образом, модель учится представлению нормального поведения этих устройств. После обучения автоэнкодера (с использованием предлагаемого подхода раздельного обучения) он оценивается на сетевых трассировках, содержащих легитимные и атакующие образцы.

Ошибка реконструкции, рассчитываемая как расхождение между входными данными и результатами их воспроизведения на выходном слое, выступает в качестве метрики для оценки степени аномальности вновь поступающих данных. Если новые сетевые образцы соответствуют тому же распределению, что и обучающие данные, величина ошибки реконструкции окажется незначительной. В то же время предполагается, что образцы, связанные с атаками, будут существенно отличаться от обученного распределения данных в результате чего значение ошибки реконструкции заметно возрастёт.

В данной работе представлена архитектура кластеризованного раздельного обучения, предназначенная для обучения неконтролируемых моделей обнаружения аномалий и вторжений в масштабных сетях, состоящих из гетерогенных устройств IoT. Ключевое преимущество предложенного подхода заключается в отсутствии необходимости контролируемой маркировки

данных, что делает его особенно пригодным для реальных развёртываний, где точная маркировка сетевого трафика зачастую невозможна. Для преодоления сложностей, связанных с применением в гетерогенных средах, архитектура включает неконтролируемый алгоритм кластеризации устройств, который анализирует параметры частично обученных моделей. Данный метод кластеризации полностью интегрирован в конвейер раздельного обучения, не требует использования внешних инструментов дактилоскопии или ручных методов кластеризации, что существенно упрощает внедрение в реальных условиях эксплуатации.

Вместе с тем выявлено, что качество кластеризации может снижаться при недостаточном объёме локальных обучающих данных на отдельных устройствах. Однако в широком диапазоне объёмов обучающих данных результаты кластеризации оставались стабильными, а потенциальные проблемы можно смягчить за счёт обеспечения достаточного количества данных перед стартом процесса. Дополнительным решением может стать многократное выполнение шага кластеризации на сервере для повышения стабильности процесса. Поскольку данный шаг не требует значительных локальных вычислений и предполагает лишь один раунд коммуникации, его ресурсные затраты остаются невысокими.

Кроме того, обучение с применением продемонстрировало более высокую скорость сходимости модели по сравнению с изолированным обучением на периферийных устройствах, особенно для тех из них, которые генерируют небольшие объёмы обучающих данных. Оценка глобальных моделей на реальных атаках показала низкие показатели ложных срабатываний и высокую эффективность обнаружения большинства атак. Хотя некоторые DoS-атаки не были корректно классифицированы как аномальные для отдельных кластеров устройств, предложенная модель успешно выявляла более скрытые вредоносные действия – например, пакеты heartbeat C&C Mirai и медленные сканирования.

При сравнении с ML-основанной сетевой IDS Kitsune выяснилось, что последняя корректно обнаруживает DoS-атаки, но ошибочно классифицирует более скрытую активность. Это наводит на мысль о целесообразности совместного использования кластеризованной модели раздельного обучения с более простой моделью, которая, к примеру, анализирует частоту пакетов в заданном временном окне для выявления объёмных атак. Также предложенный кластеризованный подход продемонстрировал превосходство над некластеризованными базовыми реализациями раздельного обучения. Обучение единой глобальной модели для всех гетерогенных устройств выявило предвзятость в пользу устройств, генерирующих больший объём обучающих данных, либо недостаточную обобщающую способность такой модели. Это подчёркивает преимущества персонализации, обеспечиваемой кластеризованным раздельным обучением, в задачах неконтролируемого обнаружения сетевых аномалий.

Заключение

Включение устройств с высокой мобильностью (например, интеллектуальных транспортных средств и БПЛА) создаёт дополнительные сложности из-за их частых переходов между беспроводными сетями с разным качеством обслуживания, что приводит к непрерывным изменениям извлекаемых сетевых признаков. Оценка и адаптация неконтролируемых методов кластеризованного обучения для сценариев с высокой мобильностью представляют собой перспективное направление будущих исследований. Если распределение данных устройства изменяется после кластеризации, но до завершения полного цикла раздельного обучения, можно применить динамические или мягкие методы кластеризации для повышения гибкости в сетях IoT с высокой мобильностью. Ещё одним направлением будущих работ является анализ первопричин аномалий в частности, различение вторжений и атак от иных факторов (например, обновлений устройств).

Наконец, следует отметить, что неконтролируемое обучение модели предполагает работу устройств в нормальных условиях, то есть на этапе обучения устройства не должны быть скомпрометированы. Данное допущение может не выполняться в некоторых враждебных сценариях. В будущих исследованиях планируется изучить, как скомпрометированные или враждебные устройства влияют на неконтролируемый этап кластеризации в предложенном методе. Такие устройства могут демонстрировать поведение, отличающееся от нормальных устройств того же кластера, и этот дрейф способен служить индикатором аномальной

активности. В подобных случаях устройство может быть помечено или исключено из процесса до его начала.

Литература

1. Буйневич, М. В. Вопросы обеспечения информационно-технологического суверенитета в Арктике в контексте актуальных проблем информационной безопасности России / М. В. Буйневич // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2014) : III Международная научно-техническая и научно-методическая конференция: сборник научных статей, Санкт-Петербург, 25–26 февраля 2014 года. – Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2014. – С. 39-47.
2. Метод обнаружения аномалий в сетевом компьютерном трафике на основе нейронной сети с использованием LSTM / А. М. Крибель, О. С. Лаута, А. В. Филин, А. С. Фень // Электросвязь. – 2021. – № 12. – С. 43-48.
3. Барсуков, И. С. Алгоритм анализа фрактальных свойств трафика для обнаружения сетевых аномалий / И. С. Барсуков, М. П. Ряполов, А. М. Бобрешов // Радиолокация, навигация, связь : Сборник трудов XXVI Международной научно-технической конференции, в 6 т., Воронеж, 29 сентября – 01 2020 года. Том 4. – Воронеж: Воронежский государственный университет, 2020. – С. 302-311.
4. Мамчев Г. В. Основы радиосвязи и телевидения. Учебное пособие для вузов. - М: Горячая линия-Телеком, 2007. - 416 с.:
5. Федорова, М. Л. Об исследовании свойства самоподобия трафика мультисервисной сети / М. Л. Федорова, Т. М. Леденева // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2010. – № 1. – С. 46-54.

DEVELOPMENT OF A NETWORK ANOMALY DETECTION DEVICE FOR HETEROGENEOUS LARGE-SCALE IOT NETWORKS

PhD, Assoc. Prof. I.A. Sorokin, Lect. A.D. Chesnokov

Nizhny Novgorod University of Engineering and Economics

This article explores the development of a clustered learning architecture for detecting network anomalies in large-scale heterogeneous Internet of Things (IoT) networks. The relevance of this research stems from the exponential growth of attack vectors and the low effectiveness of traditional signature-based protection methods in the face of malware polymorphism and high device heterogeneity. The authors propose an approach based on clustered learning that enables training unsupervised models (autoencoders) on distributed data without centralized data collection, ensuring privacy and reducing network load. A key element of the architecture is the integration of an unsupervised device clustering algorithm that analyzes the parameters of partially trained models to generate adaptive global models specific to groups of devices with similar behavior patterns.

ПРИНЦИПЫ ОЦЕНКИ УРОВНЯ МНОГОЛУЧЕВОГО СИГНАЛА

асп. Мухин Н.А., к.т.н., доц. Шахтанов С.В.

Нижегородский инженерно-экономический университет

Распространение электромагнитных волн в свободном пространстве в условиях неограниченных сред носит прямолинейный и однонаправленный характер. Однако в реальных условиях прямолинейность распространения нарушается в зависимости от местных условий и частотного диапазона. В длинноволновых и коротковолновых участках диапазона наблюдаются искривления траектории распространения за счёт рефракции и разнообразных отражений от различных слоёв ионосферы. В более высокочастотных диапазонах, условно от 100 мегагерц, когда длина волны становится соразмерной с размерами местных препятствий, начинают оказывать значимое влияние более сложные механизмы распространения. Часть фронта волны распространяется напрямую в пределах прямой видимости, другие составляющие сигнала отражаются от крупных фрагментов рельефа и фасадов зданий, огибают препятствия за счёт явлений дифракции, рассеиваются на атмосферных неоднородностях и на малогабаритных фрагментах рельефа, примером которых является лесистая местность. В силу названных причин, в реальных условиях использования высокочастотного участка диапазона, имеет место многолучевой характер распространения волны, когда в точку приёма приходят одновременно несколько сигналов от одного источника с разной фазой, амплитудой и при наличии быстрых и медленных замираний уровня в ходе обмена данными. Анализ вариантов развития событий крайне необходим на этапе проектирования беспроводных систем связи, особенно с учётом перспектив внедрения сетей 5G и 6G.

Уравнение Фрииса: базовый эталон для свободного пространства

В середине 40-х годов 20 века Гарольд Трап Фриис [2] предложил простую формулу для расчёта мощности сигнала в условиях свободного пространства. Согласно его уравнению, принимаемая мощность сигнала пропорциональна произведению коэффициентов усиления антенн, квадрату длины волны, и обратно пропорциональна квадрату расстояния между антеннами:

$$P_r = P_t G_t G_r (\frac{\lambda}{4\pi d})^2 \quad (1)$$

где P_r — мощность сигнала на входе приёмной антенны; P_t — мощность, излучаемая передающей антенной, G_t и G_r — коэффициенты усиления передающей и приёмной антенн соответственно, λ — длина волны электромагнитного излучения, d — расстояние между фазовыми центрами передающей и приёмной антенн;

Формула была выведена во времена массового освоения УКВ диапазона для расчёта энергетического бюджета радиорелейных линий связи. Эта модель остаётся отправной точкой для проектирования и анализа линий радиосвязи, но её ограничения становятся очевидными в реальных условиях. Если говорить только о радиорелейных линиях связи, особенностью которых являются узконаправленные антенны, то при расчёте радиорелейной трассы в общем случае достаточно определить границы первой зоны Френеля (рис. 1).

Радиус первой зоны Френеля [1] в направлении перпендикулярном направлению распространения волны можно определить из выражения:

$$b_1 = \sqrt{\frac{l_1 l_2 \lambda}{l_1 + l_2}}, \quad (2)$$

где l_1 – расстояние от передающей антенны до точки, где определяются границы зоны Френеля, l_2 – расстояние от этой точки до приёмника.

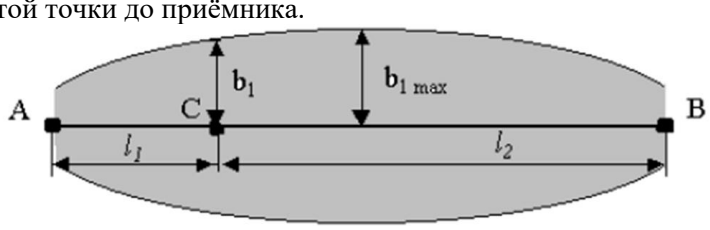


Рисунок 1. Область, существенная для распространения волны
A – передатчик, *B* – приёмник, b_1 – радиус первой зоны Френеля

Открытость интервала радиотрассы зависит от взаимного расположения антенн и их места установки, при этом попадание местных топографических элементов в зону прохождения

сигнала можно компенсировать высотой подвески антенн. Именно поэтому формула Фрииса не учитывает многолучёвости распространения сигнала, поскольку из-за узких диаграмм направленности априори выбираются условия, где нет отражений от местных элементов архитектуры и деревьев.

Реальное распространение сигнала без предварительного учёта характеристик антенн и их диаграмм направленности происходит в условиях наличия местных форм окружающего пространства в виде зданий, деревьев, особенностей рельефа и погодных аномалий, вносящих поправки, которые могут в десятки и сотни раз превосходить базовые потери по Фриису (рис. 2). Имеющие место отклонения от этого «эталонного» расчёта позволяют количественно оценить влияние окружающей среды на распространение сигнала.

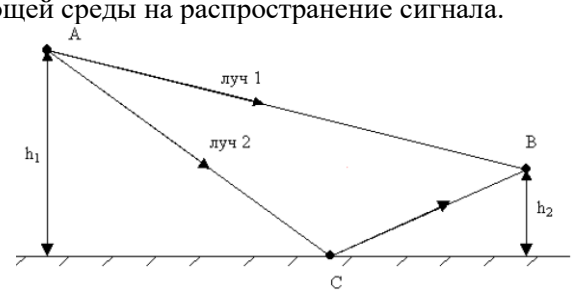


Рисунок 2. Пути прохождения лучей от пункта передачи (точка A), до пункта приёма (точка B), C – точка отражения луча 2 от поверхности земли или местных форм

Приведём пример расчёта с использованием формулы Фрииса без учёта влияния многолучёвости для конкретных исходных данных, типичных для городских условий:

Частота: f=900МГц

Мощность передатчика: P_t=1Вт =30дБм

Коэффициенты усиления антенн: G_t=G_r=1 (изотропные антенны, 0 дБ)

Расстояние: d =1км =1000м

Определение длины волны:

$$\lambda = \frac{c}{f} = \frac{3{\cdot}10^8 \text{ м/с}}{900{\cdot}10^6 \text{ Гц}} \approx 0{,}333 \text{ м.} \quad (3)$$

где c — скорость света в вакууме, f — частота сигнала;

Отсюда по формуле Фрииса:

$$P_r = 1 \cdot 1 \cdot 1 \cdot \left(\frac{0{,}333}{4\pi{\cdot}1000}\right)^2 \approx \left(\frac{0{,}333}{12566{,}4}\right)^2 \approx (2{,}65 \cdot 10^{-5})^2 \approx 7{,}02 \cdot 10^{-10} \text{ Вт.} \quad (4)$$

Переведём в децибелы относительно 1 мВт (дБм):

$$P_r(\text{дБм}) = 10\lg(7{,}02 \cdot 10^{-7}) \approx -61{,}5 \text{ дБм.} \quad (5)$$

Таким образом, в условиях свободного пространства на расстоянии 1 км от передатчика приёмник получит сигнал мощностью около -61,5 дБм.

Модель Окамура–Хата: учёт городской среды

В современных условиях широкого распространения сотовой связи наиболее востребованы расчёты зон покрытия в условиях городской местности различной плотности застройки. Это наиболее характерный случай многолучевого распространения сигнала. Натурные испытания на этапе проектирования строго индивидуальны от проекта к проекту относительно каждого варианта установки базовых станций, поэтому затратны и не эффективны. Решением проблемы является использование усреднённых эмпирических моделей.

Одной из наиболее удачных и широко используемых в качестве исходных данных моделей, является модель японского исследователя Юкио Окамура[3], проводившего масштабные измерения в Токио в 1960-х годах. На основе этих измерений японцем Масахару Хата в 1980 году была разработана аналитическая модель для прогнозирования потерь на пути передачи радиосигналов в условиях городской, пригородной и сельской местностей. Полученная модель учитывает частоты в диапазоне 150–1500 МГц, высоту антенн базовой и подвижной станций, расстояние между ними и тип застройки.

Для среднего города формула медианных потерь сигнала имеет вид:

$$L_{50} = 69{,}55 + 26{,}16\lg f - 13{,}82\lg h_b - a(h_m) + (44{,}9 - 6{,}55\lg h_b)\lg d \quad (6)$$

где L_{50} – медианные потери на трассе (в децибелах), то есть уровень ослабления сигнала, относительно которого в реальных условиях измерения в 50 % случаев потери будут ниже, а в 50 % случаев — выше, f — частота в мегагерцах, h_b и h_m — высоты антенн в метрах, d — расстояние в километрах, $a(h_m)$ — поправка, зависящая от высоты приёмной антенны,. Индекс «50» указывает на 50-й процентиль статистического распределения потерь пути.

Для малых и средних городов при f=900 МГц используется приближение:

$$a(h_m) = (1{,}1 \lg f - 0{,}7)h_m - (1{,}56 \lg f - 0{,}8) \quad (7)$$

Дополнив исходные данные для формулы Фрииса параметрами высот установки антенн: высота базовой станции: h_b=30м, высота мобильной станции: h_m=1,5м, рассчитаем затухание сигнала в городских условиях при наличии изменения интерференционной картины многолучёвости за счёт имеющего место передвижение абонентского устройства.

Расчёт поправки $a(h_m)$ из формулы 7: $\lg f = \lg 900 \approx 2{,}954$,

$$a(h_m) = (1{,}1 \cdot 2{,}954 - 0{,}7) \cdot 1{,}5 - (1{,}56 \cdot 2{,}954 - 0{,}8) \approx (3{,}249 - 0{,}7) \cdot 1{,}5 - $$
$$-(4{,}608 - 0{,}8) \approx 2{,}549 \cdot 1{,}5 - 3{,}808 \approx 3{,}824 - 3{,}808 \approx 0{,}016 \text{ дБ}.$$

Итоговой расчёт потерь мощности сигнал из формулы 6:

$$L_{50} = 69{,}55 + 26{,}16 \cdot 2{,}954 - 13{,}82 \cdot \lg 30 - 0{,}016 + (44{,}9 - 6{,}55 \cdot \lg 30) \cdot \lg 1.$$

при lg1=0 последнее слагаемое равно нулю, lg30≈1,477:

$$L_{50} \approx 69{,}55 + 77{,}28 - 13{,}82 \cdot 1{,}477 - 0{,}016 \approx 69{,}55 + 77{,}28 - 20{,}42 - 0{,}016 \approx 126{,}4 \text{ дБ}.$$

Таким образом, медианные потери мощности сигнала составляют 126,4 дБ. При мощности передатчика 30 дБм (1 Вт) и единичных коэффициентах усиления антенн мощность сигнала на входе приёмной антенны P_t составит:

$$P_t = 30 - 126{,}4 = -96{,}4 \text{ дБм}. \quad (8)$$

Это значение значительно ниже, чем по модели Фрииса (–61,5 дБм), что наглядно демонстрирует влияние городской застройки: дополнительные потери составляют почти 35 дБ.

Квадратичная формула Введенского

В отечественной радиотехнической школе значительное место занимает подход, предложенный одним из первых исследователей распространения сигналов УКВ диапазона в свободном пространстве - Борисом Николаевичем Введенским в 20-х годах 20 века [4]. Его эмпирическая формула учитывает влияние проводящей земной поверхности на распространение радиоволн в условиях, близких к свободному пространству, но с обязательным учётом отражённой волны от поверхности Земли. Для практических расчётов напряжённости поля в условиях прямой видимости Введенский предложил соотношение:

$$E = \frac{4\pi h_t h_r \sqrt{30 P_t G_t}}{\lambda d^2}. \quad (9)$$

где E — напряжённость электрического поля в точке приёма (В/м); Pt — мощность передатчика (Вт); G_t — коэффициент усиления передающей антенны; h_t и h_r — высоты передающей и приёмной антенн над поверхностью Земли (м); λ — длина волны (м); d — расстояние между антеннами (м).

Для удобства инженерных расчётов часто используется логарифмическая форма, выражающая потери пути L в децибелах:

$$L = 40 \lg d - 20. \quad (10)$$

где d задаётся в километрах, высоты антенн — в метрах, частота f — в мегагерцах, а результат L получается в децибелах.

Применим эту формулу к ранее заданным условиям городской застройки:

- Частота: f=900МГц
- Мощность передатчика: P_t=1Вт=30дБм
- Высота базовой станции: h_t =30м
- Высота мобильной станции: h_r =1,5м
- Расстояние: d=1км
- Коэффициенты усиления антенн: G_t=G_r =1 (0 дБ)

Подставим значения в логарифмическую формулу (8):

$$\lg d = \lg 1 = 0,$$
$$h_t h_r = 30 \cdot 1{,}5 = 45 \text{м}^2,$$

$$\lg(h_t h_r) = \lg 45 \approx 1{,}653,$$
$$\lg f = \lg 900 \approx 2{,}954.$$

Тогда потери пути составят:

$$L = 40 \cdot 0 - 20 \cdot 1{,}653 + 20 \cdot 2{,}954 - 28 \approx 0 - 33{,}06 + 59{,}08 - 28 \approx -2{,}0 \text{ дБ}.$$

Отрицательное значение потерь формально указывает на то, что в условиях идеальной отражающей поверхности и при данных высотах антенн интерференция прямой и отражённой волн может приводить к усилению сигнала относительно свободного пространства. Однако на практике реальная земная поверхность обладает конечной проводимостью и неоднородностью рельефа, поэтому для корректного применения формулы Введенского вводится поправочный коэффициент ослабления отражённой волны $K_{осл}$, зависящий от типа местности (водная поверхность, влажная почва, сухая земля, городская застройка). Для типичных условий сухой почвы принимают $K_{осл} \approx 0{,}6$ (–4,4 дБ), для городской среды — $K_{осл} \approx 0{,}3$ (–10,5 дБ).

С учётом поправки для городских условий (−10,5дБ) суммарные потери составят:

$$L_{корр} = -2{,}0 + 10{,}5 = 8{,}5 \text{ дБ}.$$

Тогда мощность на входе приёмника:

$$P_r = P_t - L_{корр} = 30 - 8{,}5 = 21{,}5 \text{ дБм}.$$

Этот результат явно противоречит реальности и демонстрирует ограничения применимости уравнения Введенского для ультракоротких волн в городских условиях. Исторически формула изначально разрабатывалась для средних и длинных волн (сотни килогерц) наиболее востребованных в тот период времени для радиосвязи при условиях, что радиус кривизны Земли и длина волны сопоставимы, а отражение от поверхности играет определяющую роль. На частотах 900 МГц и выше основным фактором становятся отражения от зданий, дифракция на границах препятствий и рассеивание, явление, которое модель Введенского не учитывает. Поэтому для диапазона УКВ/СВЧ в современной практике применяется модель Окамура–Хата и её варианты для различных диапазонов, специально разработанные для городской застройки.

Распределение Райса: статистика мелкомасштабных замираний

Для прогнозирования уровня сигнала в точке приёма при условии множественных отражений применимы статистические методы. Если в точку приёма приходит кроме прямого луча множество слабых рассеянных лучей, амплитуда суммарного сигнала подчиняется распределению Райса [5]. Ключевым параметром здесь выступает фактор K — отношение мощности доминирующей компоненты к аддитивной мощности рассеянных составляющих:

$$K = \frac{A^2}{2\sigma^2} \quad (11)$$

где A - амплитуда детерминированной (прямой или квазипрямой) компоненты сигнала, приходящей по линии прямой видимости или отражённой от крупного объекта,

σ^2 — дисперсия (средняя мощность) рассеянных компонентов сигнала, не имеющих доминирующего направления.

Плотность вероятности мгновенной амплитуды r в райсовском канале (канал с отражениями, но с преобладанием прямого распространения сигнала) задаётся выражением:

$$f(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2+A^2}{2\sigma^2}\right) I_0\left(\frac{rA}{\sigma^2}\right), r \geq 0 \quad (12)$$

где $f(r)$ — плотность вероятности мгновенной амплитуды сигнала, r — текущее значение амплитуды, огибающей радиосигнала в момент наблюдения, I_0 — модифицированная функция Бесселя первого рода нулевого порядка, возникающая при интегрировании распределения фаз случайных компонентов по всем направлениям (функция Бесселя от чисто мнимого аргумента. является решением дифференциального уравнения Бесселя, в котором аргумент X может быть чисто мнимым).

Используем тот же сценарий: частота 900 МГц, расстояние 1 км, городская среда. Допустим, что в данной точке приёма наблюдается частичная прямая видимость, и фактор Райса составляет K=6дБ, что соответствует линейному значению $K=10^{6/10} \approx 4$ (перевод фактора Райса из децибельной шкалы в линейную).

Общая средняя мощность сигнала равна сумме мощностей прямой и рассеянных компонент:

$$P_{avg} = A^2 + 2\sigma^2 \quad (13)$$

Из определения $K=A^2/(2\sigma^2)$ следует $A^2=2K\sigma^2$, тогда:

$$P_{avg} = 2K\sigma^2 + 2\sigma^2 = 2\sigma^2(K+1) \quad (14)$$

Из расчёта по модели Окамура–Хата следует, что средняя мощность сигнала составляет –96,4 дБм, или в линейных единицах:

$$P_{avg} = 10^{-96,4/10} \cdot 10^{-3} \approx 2,29 \cdot 10^{-13} \text{ Вт.}$$

Найдём σ^2:

$$\sigma^2 = \frac{P_{avg}}{2(K+1)} = \frac{2,29 \cdot 10^{-13}}{2(4+1)} = \frac{2,29 \cdot 10^{-13}}{10} = 2,29 \cdot 10^{-14}.$$

Тогда амплитуда прямой компоненты:

$$A = \sqrt{2K\sigma^2} = \sqrt{2 \cdot 4 \cdot 2,29 \cdot 10^{-14}} = \sqrt{1,832 \cdot 10^{-13}} \approx 4,28 \cdot 10^{-7} \text{ В.}$$

Исходя из результата, можно оценить вероятность, что мгновенная мощность сигнала упадёт ниже порогового уровня, например, –100 дБм (что соответствует амплитуде $r_{th}=\sqrt{10^{-10}}\approx10^{-5}$ В).

Однако более практично использовать функцию распределения мощности или численные методы. Важно отметить, что при K=4 глубокие замирания, например ниже –110 дБм, происходят существенно реже, чем в рэлеевском канале при K=0 [6] (модель радиоканала с многолучевым распространением, в которой отсутствует доминирующая прямая или отражённая компонента сигнала, а принимаемый сигнал формируется исключительно множеством рассеянных компонентов, приходящих из разных направлений со случайными амплитудами и фазами), где вероятность таких событий может достигать нескольких процентов.

Заключение

Таким образом, последовательное применение трёх ключевых моделей — Фрииса, Окамура–Хата, Введенского и статистического распределения Райса позволяет получить достаточно полную картину распространения радиосигнала: от идеального случая Фрииса до крупномасштабных потерь в городской среде и до статистического описания мелкомасштабных флуктуаций. В рассмотренном примере мощность сигнала падает от –61,5 дБм (Фриис) до –96,4 дБм (город по Окамура–Хата), а затем дополнительно подвергается случайным колебаниям, характерным для райсовского канала с K=4. Такой многоуровневый подход необходим для надёжного проектирования современных беспроводных систем, где каждая дополнительная потеря мощности сигнала и каждый процент вероятности замирания могут оказаться критичными.

Литература

1. Кубанов В.П., Ружников В.А., Сподобаев М.Ю., Сподобаев Ю.М. Основы теории антенн и распространения радиоволн: Учебное пособие / Под ред. В.П. Кубанова. – С.: ИНУЛ-ПГУТИ, 2016. – 258 с.: ил. ISBN 978-5-9912-0152-0.

2. Кугушев А.М., Голубева Н.С., Митрохин В.Н. Основы радиоэлектроники. Электродинамика и распространение радиоволн. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2001. – 368 с.

3. Бузова М.А. Электродинамический анализ сложных антенных комплексов на основе интегральных уравнений. Монография. – М.: Радиотехника, 2025. – 352 с. ISBN 978-5-93108-242-4.

4. Крыжановский В.Г. Техническая электродинамика. – Донецк: ДонГУ, 2003. – 116 с.

5. Мамчев Г. В. Основы радиосвязи и телевидения. Учебное пособие для вузов. - М: Горячая линия-Телеком, 2007. - 416 с.: ил. ISBN 5-93517-267-4.

6. Черенкова Е. Л., Чернышев О. В. Распространение радиоволн: Учебник для вузов связи. — М.: Радио и связь, 1984. — 272 с., ил.

7. Дворников С.В., Балыков А.А., Котов А.А. Упрощенная модель расчета потерь сигнала в радиолинии, полученная путем сравнения квадратичной формулы Введенского с существующими эмпирическими моделями // Системы управления, связи и безопасности. 2019. №2. URL: https://cyberleninka.ru/article/n/uproschennaya-model-rascheta-poter-signala-v-radiolinii-poluchennaya-putem-sravneniya-kvadratichnoy-formuly-vvedenskogo-s (дата обращения 30.01.2026г.)

PRINCIPLES OF MULTIPATH SIGNAL LEVEL ASSESSMENT

Asst. Mukhin N.A., Ph.D. Shakhtanov S.V.

Nizhny Novgorod University of Engineering and Economics

The propagation of electromagnetic waves in free space in unbounded media is rectilinear and unidirectional. However, in real-world conditions, this rectilinear propagation is disrupted depending on local conditions and the frequency range. In the long- and short-wavelength ranges, propagation trajectories are distorted due to refraction and various reflections from various ionospheric layers. In higher frequency ranges, say, from 100 megahertz, when the wavelength becomes commensurate with the size of local obstacles, more complex propagation mechanisms begin to have a significant impact. Part of the wavefront propagates directly within the line of sight, while other signal components are reflected from large terrain features and building facades, bend around obstacles due to diffraction, and are scattered by atmospheric irregularities and small terrain features, such as wooded areas. For these reasons, in real-world high-frequency applications, multipath propagation occurs, with multiple signals arriving simultaneously from a single source, each with different phase and amplitude, and with both fast and slow fading levels during data exchange. Analyzing these scenarios is essential during the design phase of wireless communication systems, especially given the prospects for 5G and 6G networks.

РЕГУЛЯРНАЯ КОНЕЧНАЯ МАРКОВСКАЯ ЦЕПЬ, ОПИСЫВАЮЩАЯ ТРЕХАЛЬТЕРНАТИВНЫЙ ПЯТИЭТАПНЫЙ ПРОЦЕСС ПРИНЯТИЯ РЕШЕНИЯ ОПЕРАТОРОМ АРМ АСУ ОБЩЕГО НАЗНАЧЕНИЯ

д.т.н., доц. Попов М.Ю.[1], к.т.н. Девятияров В.Б.[2],
к.т.н. Слободсков О.Е.[3], Попова Н.В.[1]

[1]Филиал военной академии РВСН имени Петра Великого в г. Серпухове.
[2]АО «Корпорация «Московский институт теплотехники», г. Москва.
[3]АО «Российский институт мощного радиостроения», г. Санкт-Петербург.

В статье представлена математическая модель в виде регулярной конечной марковской цепи, описывающая трехальтернативный пятиэтапный процесс принятия решения оператором автоматизированного рабочего места автоматизированной системы управления с системой поддержки принятия решения. Система поддержки принятия решения функционирует в поэтапном режиме. Представлен граф состояний и переходов и регулярная переходная матрица. Описан физический смысл состояний и аналитические выражения определяющие численных характеристики цепи.

В настоящее время большое распространение получили автоматизированные системы управления (АСУ) с поддержкой принятия решения (ППР) [1-3,5]. Оператор автоматизированного рабочего места (АРМ) такой АСУ при поступлении по каналам связи сообщения и отображении его на терминальном устройстве реализует процесс принятия решения (отработку поступившего сообщения) [2,4,5] На принятие решения оператору предоставляется определенное время, которое зависит от содержательной части сообщения. Оператор может сразу принять решение по отработке поступившего сообщения, а может дождаться предоставления системой ППР дополнительной информации и после осуществить принятие решения, потеряв при этом время. Результатом работы оператора по отработке поступившего сообщения есть три альтернативы: принято правильное решение; принято не правильное решение; решение не принято. Система ППР предоставляет дополнительную необходимую информацию оператору в четыре этапа. Каждый этап осуществляется в определенный момент времени (t_1, t_2, t_3, t_4). Таким образом, максимальное возможное количество этапов процесса принятия решения будет пять. Данные моменты времени фиксированы и не зависят от действий оператора. При получении оператором данной информации вероятности и времена достижения альтернатив естественно будут меняться. Если оператор принимает правильное решение, то предоставление дополнительной информации системой ППР прекращается. Такой процесс принятия решения оператором возможно формализировать на базе научно-методического аппарата регулярных конечных марковских цепей (РКМЦ). Данная формализация предполагает синтез графа состояний и переходов описывающего рассматриваемый процесс, он представлен на рисунке 1.

Физический смысл состояний таков:

S_0 - состояние, соответствующее поступлению сообщения с соответствующим отображением его на пульте оператора автоматизированного рабочего места АСУ. Оператор начал формирования решения по отработке поступившего сообщения;

S_1, (S_2) - на первом ЭПР оператором не принято решение (принято неправильное решение) по отработке поступившего сообщения;

S_3, S_6, (S_4, S_5) - на втором ЭПР оператором не принято решение (принято неправильное решение) по отработке поступившего сообщения;

S_7, S_9, S_{11}, S_{13}, (S_8, S_{10}, S_{12}, S_{14}) - на третьем ЭПР оператором непринято решение (принято неправильное решение) по отработке поступившего сообщения;

S_{15}, S_{17}, S_{19}, S_{21}, S_{23}, S_{25}, S_{27}, S_{29}, (S_{16}, S_{18}, S_{20}, S_{22}, S_{24}, S_{26}, S_{28}, S_{30}) - на четвертом ЭПР оператором не принято решение (принято неправильное решение) по отработке поступившего сообщения;

S_{31}, S_{32}, S_{33} - принятие оператором правильного решения, неправильного решения, непринятию оператором решения, соответственно.

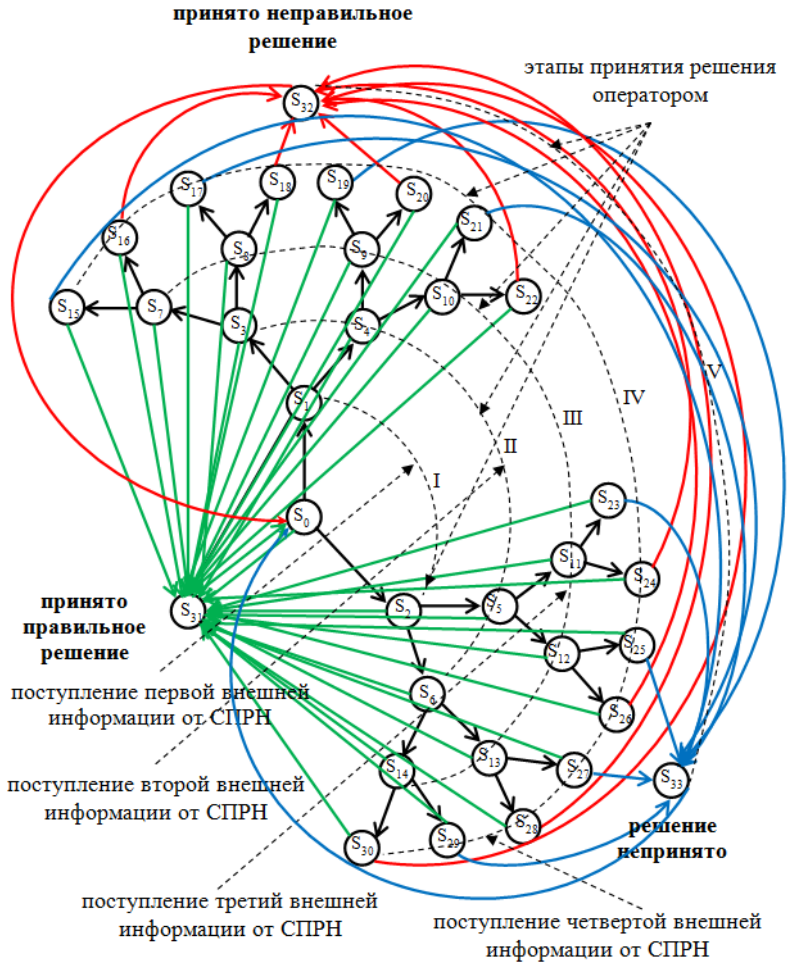


Рисунок 1 - Регулярная марковская цепь процесса принятия решения

Цепь может быть описана следующей регулярной переходной матрицей (РПМ):

$$P_{[34,34]} = \begin{Vmatrix} 0 & p_{0,1} & p_{0,2} & 0 & \dots & p_{0,31} & 0 & 0 \\ 0 & 0 & 0 & p_{1,3} & \dots & p_{1,31} & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & p_{2,31} & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & p_{3,31} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p_{31,0} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ p_{32,0} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ p_{33,0} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{Vmatrix}. \quad (1)$$

Вся информация о процессе содержаться в РПМ (1), она же формируют вероятностный уклон его поведения. Посредством численного анализа данной цепи, возможно, определить интересующие характеристики процесса принятия решения оператором, в том числе темповые.

Анализ возможностей применяемого научно-методического аппарата РКМЦ. Рассмотрим данную РКМЦ в общем виде: $i = [0; \max i]$ - шаги процесса, n - количество состояний цепи, $S_0, S_1, ..., S_{n-1}$ - состояния цепи, РПМ $P_{[n,n]} = \|p_{i,j}\|$. Отметим: «шаги процесса» ≡ «моменты времени»; «интервал времени» ≡ «продолжительность времени»; «число попаданий в состояние цепи» ≡ «время пребывания в состоянии цепи». Существуют следующие математические конструкции, через которые возможно определить *числовые характеристики* цепи.

1. **Предельная матрица** (ПМ) РМЦ.

При $i \to \infty$, ПМ есть:

$$A_{[n,n]} = P^{(i)}_{[n,n]} = \begin{Vmatrix} \alpha_{\langle n \rangle} \\ \alpha_{\langle n \rangle} \\ \dots \\ \alpha_{\langle n \rangle} \end{Vmatrix} = \begin{Vmatrix} a_0, a_1, \dots, a_{n-1} \\ a_0, a_1, \dots, a_{n-1} \\ \dots \\ a_0, a_1, \dots, a_{n-1} \end{Vmatrix} = \{a_{i,j}\}, \quad (2)$$

каждая строка $A_{[n,n]}$ представляет один и тот же вероятностный вектор: $\alpha_{\langle n \rangle} = \|a_0, a_1, \dots, a_{n-1}\|$. Отметим то, что каждая строка матрицы $P^{(i)}_{[n,n]}$ при любом *i*, содержит вероятности нахождения цепи, в том или ином состоянии, через *i* шагов процесса.

При существовании *начального распределения вероятностей состояний цепи* (вектор начальных состояний (ВНС)), описывающего «старт» процесса: $\pi_{\langle n \rangle} = P^{(0)}_{\langle n \rangle} = \|P^{(0)}_{S_0}, P^{(0)}_{S_1}, \dots P^{(0)}_{S_n}\|$, и при $i \to \infty$, имеем:

$$\pi_{\langle n \rangle} \cdot P^{(i \to \infty)}_{[n,n]} = A_{[n,n]}. \quad (3)$$

При динамике шагов процесса (i), по формируемым значениям переходных вероятностей (ПВ) предельной матрицы можно оценивать, как шаги процесса влияют на ПВ, т.е. определять *скорость сходимости* процесса описываемого РКМЦ (*«быстроту»* перехода процесса в *«стационарный»* режим).

2. Определение **предельного вектора** (ПрВ) РКМЦ.

Предельный вектор (предельное распределение, «неподвижный» вектор) при $i \to \infty$, есть:

$$\alpha_{\langle n \rangle} = \|a_0, a_1, \dots, a_{n-1}\| \equiv \|P^{(i)}_{S_0}, P^{(i)}_{S_1}, \dots, P^{(i)}_{S_n}\|, , \quad (4)$$

где $P^{(i)}_{S_0}, P^{(i)}_{S_1}, \dots, P^{(i)}_{S_n}$ - *предельные вероятности нахождения процесса в состояниях цепи*.

3. Определение **ожидаемой доли числа шагов процесса, после которых процесс будет находиться в состояниях цепи (доля длительности интервала времени, в течение которого РКМЦ будет находиться в своих состояниях**). При *большом количестве шагов процесса* (*i*) элементы $\alpha_{\langle n \rangle}$ предельного вектора $a_0, a_1, \dots, a_{n-1}$ представляют собой *ожидаемую долю числа шагов процесса, после которых процесс будет находиться в состояниях* $S_0, S_1, \dots, S_{n-1}$ *цепи*, т.е., на протяжении длительного интервала времени (*i*) цепь, например, будет примерно a_0 шагов процесса (примерная часть времени) находиться в состоянии S_0.

4. Определение **среднего времени достижения состояний цепи (математического ожидания (МО) числа шагов процесса для достижения состояний цепи)**. *Средние времена достижения состояний цепи (МО числа шагов процесса для достижения состояний цепи)* есть:

$$m_0 = \frac{1}{a_0},\ m_1 = \frac{1}{a_1},\ \dots,\ m_{n-1} = \frac{1}{a_{n-1}}. \quad (5)$$

5. Определение **фундаментальной матрицы** (ФМ) РМЦ.

С помощью ФМ вычисляется большинство характеристик, описывающих поведение РКМЦ.

ФМ есть:

$$Z_{[n,n]} = \left[I_{[n,n]} - \left(P_{[n,n]} - A_{[n,n]} \right) \right]^{-1} = \{z_{i,j}\}, \quad (6)$$

где $I_{[n,n]} = \begin{Vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{Vmatrix}$ - единичная матрица; $P_{[n,n]}$ - РПМ; $A_{[n,n]}$ - ПМ.

Элементы ФМ - *среднее количество шагов (средние количество моментов времени) проведенного процессом в* s_i *состояние цепи до перехода в* s_j ($M_j\left[\overline{y}^{(i)}_l\right]$, где j – номер

состояния, $\overline{y}_l^{(i)}$ - число шагов процесса, в которые процесс находился в состоянии s_i за первые i шагов). По элементам ФМ можно оценить в каких состояниях цепи процесс побывает в среднем чаще, в каких реже, элементы ФМ не обязаны быть неотрицательными, сумма элементов каждой строки ФМ равна 1.

Для двух любых ВНС: $\pi_{\langle n\rangle} = P^{(0)}_{\langle n\rangle} = \left\| P^{(0)}_{S_0}, P^{(0)}_{S_1}, ... P^{(0)}_{S_n} \right\|$; $\pi'_{\langle n\rangle} = P'^{(0)}_{\langle n\rangle} = \left\| P'^{(0)}_{S_0}, P'^{(0)}_{S_1}, ... P'^{(0)}_{S_n} \right\|$, имеем:

$$\left(\pi_{\langle n\rangle} - \pi'_{\langle n\rangle}\right) \cdot Z_{[n,n]} \tag{7}$$

- *разница между средним числом моментов времени (между МО числа моментов времени пребывания в S_i), в которые процесс находился в состоянии S_i за первые n шагов* (начальное положение плюс $n-1$ шаг), *сформированными с учетом разных начальных распределений*.

6. Определение **вектора предельных дисперсий** (ВПД) для числа попаданий в каждое состояние цепи, осуществляется по выражению:

$$\begin{gathered}\beta_{\langle n\rangle} = \left[\left[a_0\left(2\cdot z_{0,0} - 1 - a_0\right)\right]\left[a_1\left(2\cdot z_{1,1} - 1 - a_1\right)\right]...\left[a_{n-1}\left(2\cdot z_{n-1,n-1} - 1 - a_{n-1}\right)\right]\right] = \\ = \left[\beta_0; \beta_1; ...; \beta_{n-1}\right]\end{gathered} \tag{8}$$

- *предельные дисперсии для времен пребывания в состояниях цепи*, т.е. каждый элемент, есть предельная дисперсия числа попаданий в состояния цепи. Ими возможно оценивать точность прогноза (точность средних времен проведения в состояниях цепи) описываемого разными цепями или одной цепью, при разных ВНС.

7. Определение **матрицы средних времен первого достижения** (МСВПД ($M_{[n,n]}$)). МСВПД есть:

$$M_{[n,n]} = \left(I_{[n,n]} - Z_{[n,n]} + E_{[n,n]} \cdot Z_{dg[n,n]}\right) \cdot D_{[n,n]} = \left\{m_{i,j}\right\}, \tag{9}$$

где $I_{[n,n]}$ - единичная матрица; $Z_{[n,n]}$ - ФМ; $E_{[n,n]} = \begin{Vmatrix} 1 & 1 & ... & 1 \\ 1 & 1 & ... & 1 \\ ... & ... & ... & ... \\ 1 & 1 & ... & 1 \end{Vmatrix}$ - матрица, все элементы которой равны «1»; $Z_{dg[n,n]} = \begin{Vmatrix} z_{0,0} & 0 & ... & 0 \\ 0 & z_{1,1} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & z_{n-1,n-1} \end{Vmatrix}$ - диагональная матрица с диагональными элементами из ФМ $Z_{[n,n]}$; $D_{[n,n]} == \begin{Vmatrix} \frac{1}{a_{0,0}} & 0 & ... & 0 \\ 0 & \frac{1}{a_{1,1}} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \frac{1}{a_{n-1,n-1}} \end{Vmatrix}$ - диагональная матрица, с диагональными элементами из ПМ ($\frac{1}{a_{i,j}}$, *при* $i = j$).

Элементы МСВПД - математическое ожидание времени за которые процесс из S_i *впервые* попадет в S_j (среднее количество шагов процесса перехода из S_i в S_j) - *время первого*

достижения. По МСВПД можно оценивать МО времен первого достижения состояния из S_i в S_j (среднее время первого достижения любого состояния, из любого другого).

8. Определение **средних времен достижения отдельных состояний цепи.**

При заданном ВНС ($\pi_{\langle n \rangle}$), *средние времена достижения отдельных состояний цепи* можно определить как $\pi_{\langle n \rangle} \cdot M_{[n,n]}$, элементы данного вектора - средние времена достижения отдельных состояний цепи ($S_0, S_1, ..., S_{n-1}$).

Для двух любых ВНС $\pi_{\langle n \rangle}$ и $\pi'_{\langle n \rangle}$ имеем:

$$\left(\pi_{\langle n \rangle} - \pi'_{\langle n \rangle}\right) \cdot \left(I_{[n,n]} - Z_{[n,n]}\right) \cdot D_{[n,n]}, \tag{10}$$

- разница между *средними количествами числа шагов процесса, за которые цепь впервые попадет в* S_j *из начального состояния* S_0 (разница между средними временами достижения), при разных соответствующих векторах начальных состояний.

9. Определение **матрицы средних** (МС) осуществляется так:

$$\overline{M}_{[n,n]} = M_{[n,n]} - D_{[n,n]}. \tag{11}$$

Элементы МС - среднее количество шагов процесса до перехода из S_i в S_j (при $i \neq j$). Данной матрицы достаточно для описания цепи, она позволяет естественным способом выделять *минимальное число определяющих параметров* процесса.

10. Определение **матрицы дисперсий времен первого достижения** (МДВПД) таково:

$$M2_{[n,n]} = W_{[n,n]} - M_{sq[n,n]}, \tag{12}$$

где $M_{sq[n,n]} == \begin{Vmatrix} \left(M_{0,0}\right)^2 & \left(M_{0,1}\right)^2 & ... & \left(M_{0,n-1}\right)^2 \\ \left(M_{1,0}\right)^2 & \left(M_{1,1}\right)^2 & ... & \left(M_{1,n-1}\right)^2 \\ ... & ... & ... & ... \\ \left(M_{n-1,0}\right)^2 & \left(M_{n-1,1}\right)^2 & ... & \left(M_{n-1,n-1}\right)^2 \end{Vmatrix}$ - матрица, соответствующие элементы которой есть квадраты соответствующих элементов МСВПД $M_{[n,n]}$,

$W_{[n,n]} = M_{[n,n]} \cdot \left(2 \cdot Z_{dg[n,n]} \cdot D_{[n,n]} - I_{[n,n]}\right) + 2 \cdot \left(ZM_{[n,n]} - E_{[n,n]} \cdot ZM_{dg[n,n]}\right)$, где $ZM_{[n,n]} = Z_{[n,n]} \cdot M_{[n,n]}$; $ZM_{dg[n,n]}$ - диагональная матрица с диагональными элементами из $ZM_{[n,n]}$, остальные нулевые. Элементы МДВПД - дисперсии МО (времен первого достижения) за которые процесс из S_i *впервые* попадет в S_j, (*дисперсии моментов первого достижения*).

11. Определение **матрицы дисперсий** (МД) (**матрица ковариаций**). МД есть:

$$C_{[n,n]} = \left\{c_{i,j}\right\}, \tag{13}$$

где $c_{0,0} = a_0 \cdot z_{0,0} + a_0 \cdot z_{0,0} - a_0 \cdot 1 - a_0 \cdot a_0$; $c_{0,1} = a_0 \cdot z_{0,1} + a_1 \cdot z_{1,0} - a_0 \cdot 0 - a_0 \cdot a_1$;

$c_{0,2} = a_0 \cdot z_{0,2} + a_2 \cdot z_{2,0} - a_0 \cdot 0 - a_0 \cdot a_2$;...; $c_{i,j} = a_i \cdot z_{i,j} + a_i \cdot z_{j,i} - a_i \cdot \delta_{i,j} - a_i \cdot a_j$;

$\delta_{i,j} = \begin{cases} 1 \text{ при } i = j, \\ 0 \text{ при } i \neq j \end{cases}$.

12. Определение **вектора предельных дисперсий** (ВПД) для чисел попаданий в каждое состояние (предельные дисперсии для времен пребывания в состояниях цепи) (через элементы МД). ВПД есть:

$$\beta 2_{\langle n \rangle} = \left[c_{0,0}; c_{1,1}; ...; c_{n-1,n-1}\right] = \left[\beta 2_0; \beta 2_1; ...; \beta 2_{n-1}\right]. \tag{14}$$

где $c_{i,j}$ - элементы главной диагонали МД ($i = j$).

Каждый элемент ВПД есть *предельные дисперсии для числа попаданий (времен пребывания) в состояния цепи*. Отметим, выражения (8) и (14) формируют один и тоже вектор, т.е.: $\beta_{\langle n\rangle} = \beta 2_{\langle n\rangle}$.

13. Определение **предельной корреляции** (ПК) для чисел попадания в состояния S_i и S_j цепи. ПК для состояний РКМЦ S_i и S_j составит:

$$\frac{c_{i,j}}{\sqrt{c_{i,i}\cdot c_{j,j}}}, \tag{15}$$

при $i = j$ предельная корреляция обращается в «0». Предельные корреляции характеризуют степень влияния, в среднем, времени нахождения процесса в одном состоянии, на время нахождения процесса в другом состоянии (степень тесноты зависимости между случайными величинами). Предельные корреляции между временами пребывания в двух различных состояниях чаще всего будут отрицательны, так как, чем больше времени процесс находится в одном состоянии, тем меньше времени он проводит в другом.

14. Определение **плотности вероятности случайной величины** (центральная предельная теорема для марковских цепей). Вероятность попадания случайной величины $\frac{y_j^{(n)} - i\cdot a_j}{\sqrt{i\cdot b_j}}$ в интервал $[r,s]: r < s$ при числе шагов цепи $i \to \infty$ есть:

$$P\left(r < \frac{y_j^{(n)} - i\cdot a_j}{\sqrt{i\cdot b_j}} < s\right) \xrightarrow[i\to\infty]{} \frac{1}{\sqrt{2\cdot\pi}}\cdot\int_r^s e^{-\frac{x^2}{2}}\partial x, \tag{16}$$

где $y_j^{(n)}$ - число попаданий в состояние S_j цепи за первые i шагов; a_j - значение предельного вектора $\alpha_{\langle n\rangle}$ соответствующее состоянию S_j цепи; β_j - значение вектора предельных дисперсий $\beta_{\langle n\rangle}$ соответствующее состоянию S_j цепи; r, s - любые числа, такие, что $r < s$.

Так же необходимо отметить, что на базе подхода *расширения марковской цепи*, возможно сконструировать *расширенную цепь*, которая даст более «тонкую» информацию о изучаемом процессе, так же аппарат РКМЦ позволят, посредством выражений:

$$\alpha_{\langle 3\rangle} = (c-1)\cdot\left(\overline{M}_{[n,n]}^{-1}\cdot\xi_{\langle 3\rangle}\right)^T, \tag{17}$$

$$P_{[n,n]} = I_{[n,n]} + \left(D_{[n,n]} - E_{[n,n]}\right)\cdot\left(\overline{M}_{[n,n]}\right)^{-1}, \tag{18}$$

решать задач не только анализа, но и синтеза (т.е. решать *обратные* задачи). Где $\xi_{\langle 3\rangle}$ - вектор-столбец, все элементы которого равны 1, а c – постоянная.

Представленные выражения, позволяют разносторонним образом оценивать процесс принятия решения оператором, по средствам, например, следующих числовых характеристик: - средних долей времени пребывания в состояниях S_{31} (принято правильное решение), S_{32} (принято неправильное решение), S_{33} (решение оператором непринято) за i шагов процесса; - средних времен достижения состояний S_{31}, S_{32} и S_{33} (элементы m_i); - средних времен пребывания в состояниях до перехода в другие состояния за i шагов процесса (элементы $Z_{[n,n]}$); - предельных дисперсий времен пребывания в S_{31}, S_{32} и S_{33} (элементы $\beta_{\langle n\rangle} \equiv \beta 2_{\langle 3\rangle}$); - средних времен первого достижения S_{31}, S_{32} и S_{33} (элементы $M_{[n,n]}$); - дисперсий средних времен первого достижения S_{31}, S_{32} и S_{33} (элементы $M2_{[n,n]}$); - предельных корреляций времен пребывания в состояниях процесса (по элементам $C_{[n,n]}$).

Литература

1. Полумарковская цепь, описывающая действия оператора системы обеспечения вызова экстренных оперативных служб по единому номеру «112» [Текст] / М. Ю. Попов, Н. В. Попова, Д. А. Киреев, Д. Д. Лащев, А. П. Ильин // Доклады 25-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2023); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М., 2023. - Вып. XXV.,– С. 33–38. – ISBN 978-5-905278-53-2.

2. Методика определения ВВХ, ВХ и критических этапов процесса работы оператора АСУ техническим процессом [Текст] / М. Ю. Попов, Л. Н. Косарева, А. В. Подлегаев, Н. В. Попова, Ю. А. Перивалова // Междун. конф. «Радиоэлектронные устройства и системы для инфокоммуникационных технологий» (RES-2013) ; Доклады ; Серия: науч. конф. посвящ. Дню радио / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М. : ООО «Информпресс-94», 2013. – Вып. LXVIII. – С. 72–76.

3. Математическое моделирование процесса доведения сообщения в радиосети без обратной связи с повторениями и накоплением информации [Текст] / В. А. Цимбал, М. Ю. Попов, М. Ю. Дробышев // Информационные технологии в проектировании и производстве: науч.–техн. жур. – Москва, 2010. - №3 – С. 78–83.

4. Патент на изобретение 2665226 Российская Федерация. Мажоритарный элемент «5 и более из 9» [текст] / В. А. Цимбал, М. Ю. Попов, А. С. Винокуров, В. Б. Девятияров, А. М. Деркач, А. Н. Кривоногов, М. А. Лягин, Н. В. Попова (RU); - заявитель МОУ «Институт инженерной физики». – заявка № 2016101771; заявл. 21.01.2016; опубл. 28.08.2018, бюл. № 25. – 4 с.: ил.

5. Конечная полумарковская цепь, описывающая процесс отработки оператором автоматизированного рабочего места АСУ поступившего по радиоканалу сообщения [Текст] / М. Ю. Попов, В. В. Илющенко, А. М. Пономарев, П. К. Калач, Д. А. Киреев // Доклады 24-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2022); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М. : Вып. XXIV.,– С. 21–25. – ISBN 978-5-905278-49-5.

THE REGULAR FINAL MARKOVSKY CHAIN DESCRIBING THREE-ALTERNATIVE FIVE-STEP PROCESS OF DECISION-MAKING BY THE OPERATOR OF THE AUTOMATED WORKPLACE OF THE AUTOMATED CONTROL SYSTEM GENERAL IT IS APPOINTED

Dr.Sci.Tech. Popov M.Yu.[1], Cand.Tech.Sci. Devyatiyarov V.B.[2],
Cand.Tech.Sci. Slobodskov O.E.[3], Popova N.V.[1]

[1]Branch of military academy of Rocket strategic forces of a name of Peter the Great (c. Serpukhov, Moscow region).
[2]Joint-stock company corporation «Moscow institute of heat engineering».
[3]The Russian institute of a powerful radio structure, St.-Petersburg.

In article the mathematical model in the form of regular final markovsky is presented a chain, describing three-alternative five-step process of decision-making by the operator of the automated workplace of the automated control system with system of support of decision-making. the system of support of decision-making functions in a stage-by-stage mode. the state graph and transitions and a regular transitive matrix is presented. the physical sense of conditions and analytical expressions chains defining numerical characteristic is described.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СОВМЕСТНОЙ РАБОТЫ ОПЕРАТОРОВ АРМ АСУ ОБЩЕГО НАЗНАЧЕНИЯ ПО ОТРАБОТКЕ ПОСТУПИВШЕГО СООБЩЕНИЯ

д.т.н., доц. Попов М.Ю.[1], к.т.н. Девятияров В.Б.[2], к.т.н. Слободсков О.Е.[3]

[1]Филиал военной академии РВСН имени Петра Великого в г. Серпухове.
[2]АО «Корпорация «Московский институт теплотехники», г. Москва.
[3]АО «Российский институт мощного радиостроения», г. Санкт-Петербург.

В статье представлена математическая модель процесса совместной отработки поступившего сообщения двумя операторами автоматизированного рабочего места автоматизированной системы управления общего назначения. Модель создана на базе регулярной конечной марковской цепи с непрерывным временем. Представлен граф состояний и переходов, описаны состояния процесса и переходные вероятности. Приведен анализ характеристик регулярной конечной марковской цепи.

Целью совместной работы операторов является обеспечение эффективного выполнения общей задачи путем распределения обязанностей, действий и обмена информацией. Пусть участниками такого процесса являются: оператор 1 - в выполняет ведущую роль или специализируется на одной части задачи; оператор 2 - выполняет поддерживающую роль или специализируется на другой части задачи.

При поступлении по радиоканалу управления сообщения в аппаратуру передачи данных автоматизированной системы управления (АСУ) объекта управления осуществляется соответствующее отображения его на пульте автоматизированного рабочего места (АРМ) оператора (при условии положительного выполнения алгоритма повышения достоверности). При этом операторы обязаны выполнить определенную последовательность действий на пульте управления (предписанную инструкциями), направленную на отработку поступившего сообщения, а после его отработки, возобновить анализ пульта АРМ на предмет вновь поступившего сообщения [1, 2, 4, 5]. В общем случае операторы, при работе на пульте управления АРМ АСУ общего назначения [3], могут совершать «грубые» и «не грубые» ошибки. При этом, грубая ошибка оператора является той, которая возвращает оператора, по средствам алгоритмического обеспечения АСУ, к фазе обдумывания (анализа) поступившего сообщения и формирования действий по его отработке. Не грубая ошибка оператора является той, которая, по средствам алгоритмического обеспечения АСУ, возвращает оператора к началу отработки поступившего сообщения (не возвращает оператора к фазе обдумывания (анализа) поступившего сообщения).

Процесс совместной работы операторов возможно формализировать на базе конечной марковской цепи с непрерывным временем. Структура соответствующего графа состояний и переходов (ГСП) цепи представлена на рисунке 1.

Состояния графа таковы:

S_0 - состояние, соответствующее наличию сообщения в радиоканале управления АСУ;

S_1 (S_2) - состояние, соответствующее поступлению сообщения на пульт оператора 1 (2) АРМ, начало формирования действий по его отработке;

S_3 (S_4) - состояние, соответствующее началу отработки сформированных не правильных действий оператором 1 (2) по отработке поступившего сообщения;

S_5 (S_6) - состояние, соответствующее началу отработки сформированных правильных действий оператором 1 (2) поступившего сообщения;

S_7 (S_8) - состояние, соответствующие выполнению оператором 1 (2) неправильно сформированных действий по отработке поступившего сообщения;

S_9 (S_{10}) – состояние, соответствующие выполнению оператором 1 (2) правильно сформированных действий по отработке поступившего сообщения.

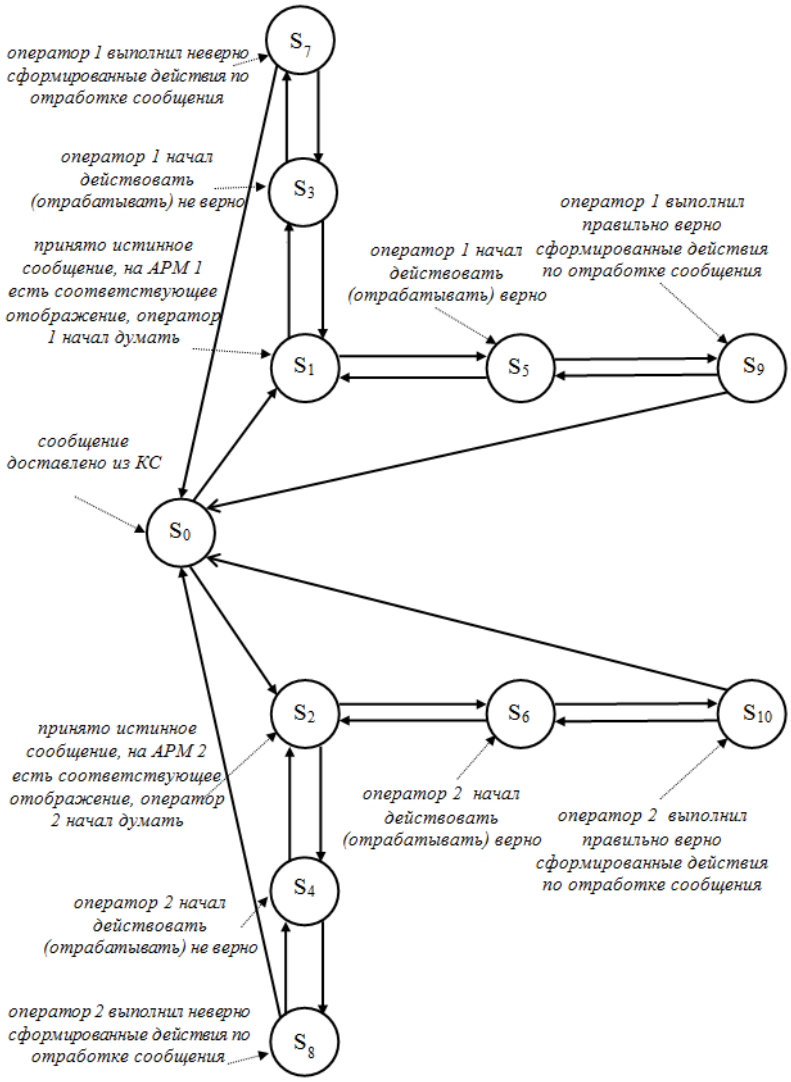


Рисунок 1 – Граф состояний и переходов марковской цепи с непрерывным временем, описывающей процесс работы операторов АРМ АСУ

Физический смысл переходных вероятностей цепи таков:

- $p_{0,1}$ ($p_{0,2}$) - вероятность поступления сообщения на пульт управления АРМ первому (второму) оператору;
- $p_{1,5}$ ($p_{2,6}$) - вероятность сформирования оператором 1 (2) правильных действий по отработке поступившего сообщения;
- $p_{5,9}$ ($p_{6,10}$) - вероятность правильного выполнения оператором 1 (2) правильно сформированных действий по отработке поступившего сообщения, с учетом возможного совершения им «не грубой» ошибки - выполнение задачи;
- $p_{5,1}$ ($p_{6,2}$) - вероятность совершения оператором 1 (2) «грубой» ошибки при выполнении правильно сформированных действий по отработке поступившего сообщения - возврат к фазе обдумывания (анализа);
- $p_{9,5}$ ($p_{10,6}$) - вероятность совершения оператором 1 (2) ошибки («не грубой») при выполнении правильно сформированных действий по отработке поступившего сообщения;
- $p_{9,0}$ ($p_{10,0}$) - вероятность возвращения оператора 1 (2) к фазе «распределения» сообщения между операторами;
- $p_{1,3}$ ($p_{2,4}$) - вероятность сформирования оператором 1 (2) не правильных действий по отработке поступившего сообщения;
- $p_{3,1}$ ($p_{4,2}$) - вероятность выполнения оператором 1 (2) некоторой последовательности неправильных действий на пульте управления АРМ по отработке поступившего по сообщения, повлекших за собой «блокировку» действий и возврат в фазу обдумывания (анализа) либо

осознание оператором 1 (2) факта выполнения им неправильных действий и самостоятельный возврат к фазе обдумывания (анализа);

- $p_{3,7}$ ($p_{4,8}$) - вероятность выполнения оператором 1 (2) некоторой последовательности неправильных действий на пульте управления АРМ по отработке поступившего по сообщения, не повлекших за собой «блокировку» действий и возврата к фазе обдумывания (анализа) и не осознание оператором 1 (2) факта выполнения им неправильных действий;

- $p_{7,3}$ ($p_{8,4}$) - вероятность совершения оператором 1 (2) «не грубой» ошибки на пульте управления АРМ при выполнении неправильно сформированных действий по отработке поступившего по сообщения;

- $p_{7,0}$ ($p_{8,0}$) - вероятность выполнения оператором 1 (2) некоторой последовательности неправильных действий на пульте управления АРМ по отработке поступившего сообщения, повлекших за собой автоматическую «блокировку» действий АСУ и возврата к фазе «распределения» сообщения между операторами.

Матрица переходных вероятностей (МПВ) такова:

$$P_{[11,11]}=\begin{Vmatrix} 0 & p_{0,1} & p_{0,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{1,3} & 0 & p_{1,5} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{2,4} & 0 & p_{2,6} & 0 & 0 & 0 & 0 \\ 0 & p_{3,1} & 0 & 0 & 0 & 0 & 0 & p_{3,7} & 0 & 0 & 0 \\ 0 & 0 & p_{4,2} & 0 & 0 & 0 & 0 & 0 & p_{4,8} & 0 & 0 \\ 0 & p_{5,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{5,9} & 0 \\ 0 & 0 & p_{6,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{6,10} \\ p_{7,0} & 0 & 0 & p_{7,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_{8,0} & 0 & 0 & 0 & p_{8,4} & 0 & 0 & 0 & 0 & 0 & 0 \\ p_{9,0} & 0 & 0 & 0 & 0 & p_{9,5} & 0 & 0 & 0 & 0 & 0 \\ p_{10,0} & 0 & 0 & 0 & 0 & 0 & p_{10,6} & 0 & 0 & 0 & 0 \end{Vmatrix}. \quad (1)$$

Таким образом, в терминах условных вероятностей МПВ, описан марковский процесс с непрерывным временем, формализирующий действия операторов АРМ, по отработке поступившего сообщения. МПВ формирует вероятностные действия операторов по средствам соответствующих переходных вероятностей. На базе анализа характеристик представленной цепи, возможно, определить ВВХ процесса отработки операторами АРМ поступившего сообщения.

Итак, марковский процесс с непрерывным временем задан МПВ. Вектор вероятностей состояний процесса в начальный момент времени есть:

$$P_{\langle 11\rangle}^{(0)}=\left\| P_{S_0}^{(0)};P_{S_1}^{(0)};P_{S_2}^{(0)};...;P_{S_{10}}^{(0)} \right\|, \quad (2)$$

где $P_{S_0}^{(0)}$ - вероятность нахождение марковского процесса формализированного представленной КМЦ в состоянии S_0 в начальный момент времени; $P_{S_1}^{(0)}$ - вероятность нахождение марковского процесса в состоянии S_1 в начальный момент времени; $P_{S_2}^{(0)}$ - вероятность нахождение в состоянии S_2 в начальный момент времени и т.д.

Для анализа характеристик процесса необходимо, в начале, определить матрицу дифференциальных уравнений (МДУ) Колмогорова, она находится на базе МПВ, следующим образом:

$$A_{[11,11]}=\left[P_{[11,11]} \right]^T-I_{[11,11]}, \quad (3)$$

где $\left[\ \right]^T$ - операция транспонирования матрицы; $I_{[11,11]}$ - единичная матрица,

$$I_{[11,11]} = \begin{Vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{Vmatrix}. \quad (4)$$

Теперь, зная вектор вероятностей состояний процесса в начальный момент времени $P^{(0)}_{\langle 11\rangle}$ и МДУ $A_{[11,11]}$ необходимо следующим образом определить правые части системы уравнений Колмогорова:

$$A_{[11,11]} \cdot \left[P^{(0)}_{\langle 11\rangle} \right]^T, \quad (5)$$

где $\left[P^{(0)}_{\langle 11\rangle} \right]^T$ - транспонированный вектор-строка вероятностей состояний процесса в начальный момент, имеющий вид:

$$\left[P^{(0)}_{\langle 11\rangle} \right]^T = \begin{Vmatrix} P^{(0)}_{S_0} \\ P^{(0)}_{S_1} \\ P^{(0)}_{S_2} \\ \dots \\ P^{(0)}_{S_{10}} \end{Vmatrix}. \quad (6)$$

Определим ВВХ рассматриваемого процесса в терминах программной среды для решения задач Mathcad, для этого необходимо создать векторную функцию двух аргументов (скалярного и векторного), она будет являться следующей функцией дифференциального уравнения Колмогорова:

$$f(t, x) = A_{[11,11]} \cdot \left[P^{(0)}_{\langle 11\rangle} \right]^T. \quad (7)$$

Таким образом, имеется система обыкновенных дифференциальных уравнений в матричном виде, для ее решения воспользуемся следующей процедурой:

$$z_{[N,12]} = rkfixed\left[\left[P^{(0)}_{\langle 11\rangle} \right]^T, t_0, t_k, N, f \right], \quad (8)$$

где $rkfixed[\cdot]$ - процедура решения системы обыкновенных дифференциальных уравнений численным методом Рунге-Кутты с фиксированным шагом; t_0 - начальный момент времени - начальная точка расчета; t_0 - конечный момент времени - коечная точка расчета; N - количество шагов на которых численный метод находит решение.

Процедура осуществляет решение системы дифференциальных уравнений Колмогорова для условных вероятностей, задающихся МПВ. В результате применения процедуры $rkfixed[\cdot]$ формируется матрица $z_{[N,12]}$. Физический смысл ее столбцов таков:

- первый столбец это динамика значений аргумента функции - времени, изменяющейся от t_0 до t_k в зависимости от заданных шагов процесса N - $t = \left[\overline{t_0, t_k} \right]$;

- второй столбец - динамика вероятности нахождения процесса в состоянии S_0 в зависимости от времени t - $P^{(t)}_{S_0}$;

- третий столбец - динамика вероятности нахождения процесса в состоянии S_1 в зависимости от времени t - $P^{(t)}_{S_1}$;

- четвертый столбец - динамика вероятности нахождения процесса в состоянии S_2 в зависимости от времени t - $P_{S_2}^{(t)}$;

- пятый столбец - динамика вероятности нахождения процесса в состоянии S_3 в зависимости от времени t - $P_{S_3}^{(t)}$ и т.д.

Таким образом, в аналитическом виде определены ВВХ процесса отработки операторами АРМ поступившего сообщения. ВВХ при этом представляют собой динамику вероятностей интересующих исследователя состояний цепи во времени.

Литература

1. Полумарковская цепь, описывающая действия оператора системы обеспечения вызова экстренных оперативных служб по единому номеру «112» [Текст] / М. Ю. Попов, Н. В. Попова, Д. А. Киреев, Д. Д. Лащев, А. П. Ильин // Доклады 25-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2023); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М., 2023. - Вып. XXV.,– С. 33–38. – ISBN 978-5-905278-53-2.

2. Методика определения ВВХ, ВХ и критических этапов процесса работы оператора АСУ техническим процессом [Текст] / М. Ю. Попов, Л. Н. Косарева, А. В. Подлегаев, Н. В. Попова, Ю. А. Перивалова // Междун. конф. «Радиоэлектронные устройства и системы для инфокоммуникационных технологий» (RES-2013) ; Доклады ; Серия: науч. конф. посвящ. Дню радио / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М. : ООО «Информпресс-94», 2013. – Вып. LXVIII. – С. 72–76.

3. Математическое моделирование процесса доведения сообщения в радиосети без обратной связи с повторениями и накоплением информации [Текст] / В. А. Цимбал, М. Ю. Попов, М. Ю. Дробышев // Информационные технологии в проектировании и производстве: науч.–техн. жур. – Москва, 2010. - №3 – С. 78–83.

4. Патент на изобретение 2665226 Российская Федерация. Мажоритарный элемент «5 и более из 9» [текст] / В. А. Цимбал, М. Ю. Попов, А. С. Винокуров, В. Б. Девятияров, А. М. Деркач, А. Н. Кривоногов, М. А. Лягин, Н. В. Попова (RU); - заявитель МОУ «Институт инженерной физики». – заявка № 2016101771; заявл. 21.01.2016; опубл. 28.08.2018, бюл. № 25. – 4 с.: ил.

5. Конечная полумарковская цепь, описывающая процесс отработки оператором автоматизированного рабочего места АСУ поступившего по радиоканалу сообщения [Текст] / М. Ю. Попов, В. В. Илющенко, А. М. Пономарев, П. К. Калач, Д. А. Киреев // Доклады 24-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2022); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М. : Вып. XXIV.,– С. 21–25. – ISBN 978-5-905278-49-5.

MATHEMATICAL MODEL OF TEAMWORK OF OPERATORS OF THE AUTOMATED WORKPLACE OF THE AUTOMATED CONTROL SYSTEM OF THE GENERAL PURPOSE ON WORKING OFF OF THE ARRIVED MESSAGE

Dr.Sci.Tech. Popov M.Yu.[1], Cand.Tech.Sci. Devyatiyarov V.B.[2], Cand.Tech.Sci. Slobodskov O.E.[3]

[1]Branch of military academy of Rocket strategic forces of a name of Peter the Great (c. Serpukhov, Moscow region).
[2]Joint-stock company corporation «Moscow institute of heat engineering».
[3]The Russian institute of a powerful radio structure, St.-Petersburg.

In article the mathematical model of process of joint working off of the arrived message is presented by two operators of the automated workplace of the automated control system of a general purpose. The model is created on the basis of regular final markov's chains with continuous time. The state graph and transitions is presented, conditions of process and transitive probabilities are described. The analysis of characteristics regular final markov's chains is resulted.

РЕШЕНИЕ ОБРАТНОЙ ЗАДАЧИ НА БАЗЕ ПОГЛОЩАЮЩЕЙ КОНЕЧНОЙ МАРКОВСКОЙ ЦЕПИ С ПОМОЩЬЮ ЯЗЫКА ПРОГРАММИРОВАНИЯ PYTHON В СРЕДЕ VISUAL STUDIO

д.т.н., доц. Попов М.Ю.[1], к.т.н. Девятияров В.Б.[2],
к.т.н. Слободсков О.Е.[3], студ. Пономарев Д.В.[4]

[1]Филиал военной академии РВСН имени Петра Великого в г. Серпухове.
[2]АО «Корпорация «Московский институт теплотехники», г. Москва.
[3]АО «Российский институт мощного радиостроения», г. Санкт-Петербург.

В статье представлен подход по нахождению переходных вероятностей матрицы переходных вероятностей поглощающей конечной марковской цепи по заданным значениям вероятностей поглощающего состояния. Создана соответствующая расчетная программа. Программа реализована на языке программирования Python в среде Visual Studio. Программа может быть использована при решении широкого круга обратных задач по определению входных параметров процессов формализованных поглощающими конечными марковскими цепями по выходным.

По выходным параметрам процесса возможно найти входные параметры их обеспечивающие - зада синтеза. Процесс нахождения таких параметров происходит путем решения обратной задачи. В зависимости от предметной области, способов формализации, применяемого научно-методического аппарата, существует множество подходов решения таких задач. Встречаются такие задачи и в теории марковских процессов, они направлены, например, на получения параметров переходных вероятностей (ПВ) матрицы переходных вероятностей (МПВ) поглощающей конечной марковской цепи (ПКМЦ) по известной динамики вероятности поглощающего состояния. Динамика вероятности поглощающего состояния, в задачах оценки качества информационного обмена, по моделям, синтезированным на базе конечных марковских цепей, может интерпретироваться как вероятностно-временные характеристики (ВВХ) процесса [1,3,4]. В общем виде такую задачу можно решать поэтапно.

Основные этапы решения таковы:

1. Формализация физического процесса.
2. Синтез поглощающей конечной марковской цепи, описывающей процесс.
3. Задания требований по ВВХ - значений вероятностей поглощающего состояния на шагах процесса (как вариант – конечного значения вероятности поглощающего состояния на последнем шаге процесса, при этом предшествующие значения вероятности поглощающего состояния должны быть обоснованы и заданы исследователем исходя из физической сущности рассматриваемого процесса).
4. Формирование МПВ цепи в символьном виде.
5. Задания вектора вероятностей состояний цепи в начальный момент времени (вектор начальных состояний) - может задаваться как заказчиком, так и исследователем, исходя их физической сущности рассматриваемого процесса.
6. Нахождение ВВХ процесса в символьном виде, например, через решение в символьном виде уравнения Колмогорова-Чепмена (УКЧ) – формирования аналитических выражений описывающих в символьном виде ВВХ.
7. Составление системы уравнений на базе сформированных аналитических выражений ВВХ на шагах процесса (п.6) и заданных их значений (п.3).
8. По необходимости упрощения, согласно логике формирования ПВ в МПВ, полученной системы уравнений.
9. Решение системы уравнений – получения значений ПВ МПВ.
10. Проверка адекватности найденных значений ПВ - расчет по ним ВВХ в численном виде, полученные при этом ВВХ должны соответствовать заданным. Если это не так, возврат к п.3.
11. По найденным переходным вероятностям определение физических параметров исследуемого процесса.

Данный подход позволяет осуществлять синтез переходных вероятностей ПКМЦ, по заданным ВВХ процесса, по их значениям установление физических параметров исследуемого процесса. Сложностью данного подхода является необходимость использования символьных

вычислений при получении аналитических выражений ВВХ, особенно если цепь рассматривается на большом количестве шагов процесса. Очередной альтернативой решения указанной обратной задачи является ее численная реализация на базе языка базе языка программирования Python. Соответствующий пример решения будет описан ниже.

Пусть задан физический процесс доставки сообщения абоненту по радиоканалу передачи данных с квитированием. Процесс формализован в виде ПКМЦ, ее граф состояний и переходов представлен на рисунке 1. Процесс доставки сообщений рассматривается на i=9 шагах процесса.

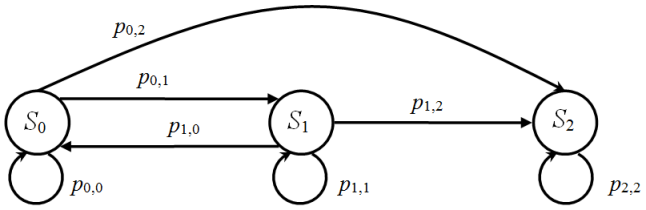


Рисунок 1 - ГСП цепи

Требуемые ВВХ заданы следующей динамикой вероятности поглощающего состояния S_2:

$$P_{S_2}^{(i=6)} = \left(P_{S_2}^{(1)}; P_{S_2}^{(2)}; P_{S_2}^{(3)}; P_{S_2}^{(4)}; P_{S_2}^{(5)}; P_{S_2}^{(6)}; P_{S_2}^{(7)}; P_{S_2}^{(8)}; P_{S_2}^{(9)}\right).$$

Вектор начальных состояний имеет вид:

$$P_{\langle 3\rangle}^{(0)} = \left(P_{S_1}^{(0)}; P_{S_2}^{(0)}; P_{S_3}^{(0)}\right).$$

МПВ цепи такова:

$$P_{[3,3]} = \begin{Vmatrix} p_{0,0} & p_{0,1} & p_{0,2} \\ p_{1,0} & p_{1,1} & p_{1,2} \\ 0 & 0 & p_{2,2} \end{Vmatrix}. \quad (1)$$

Далее представлена программа решения обратной задачи на базе языка программирования Python в среде Visual Studio.

В программе МПВ $P_{[3,3]}$ задается матрицей [P] с шестью неизвестными переходными вероятностями, которые записаны в строковом виде 'a', 'b', 'c', 'd', 'e', 'f'. Следовательно МПВ будет иметь теперь следующий вид:

$$P = \begin{Vmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{Vmatrix}, \quad (2)$$

вектор начальных состояний $P_{\langle 3\rangle}^{(0)}$ сформируем так:

P_01=['1', '0', '0'].

Матрицы в python задаются в виде следующих списков:

```
import sympy as sp
P = [['a', 'b', 'c'],
['d', 'e', 'f'],
['0', '0', '1']]

P_01 = ['1', '0', '0']
```

Вводим переменную для расчета количества неизвестных в матрице [P]:

$$elements = \left(len(P) \times len\left(P[0]\right) - len(P[-1])\right), \quad (3)$$

так как сумма переходных вероятностей каждой из строк ПКМЦ должны быть равна 1, вводим в программу следующее уравнение:

$$rr = 'a'+'+'+'b'+'+'+'c'+'-'1'. \quad (4)$$

Далее, используем следующую процедуру:

calculate_vector,

которая поэлементно перемножает значения матрицы [P] и [P_01] и записывает полученные значения в список result = [].

Получим:

```
def calculate_vector(matrix, coefficients):
    result = []
    for i in range(len(P)):
        t_i = '0'
        for j in range(len(P [0])):
            t_i = '(' + t_i + '+' + '(' + coefficients[j] + '*' + matrix[j][i] + ')' + ')'
        result.append(t_i)
    return result
result = []
```

Значения вероятностей нахождения цепи в состояниях (S_1, S_2, S_3) на первом шаге процесса ($P_{S_0}^{(1)}$; $P_{S_1}^{(1)}$; $P_{S_2}^{(1)}$) есть:

$$['\Big(\Big(\big(0+(1\times a)\big)+(0\times d)\Big)+(0\times 0)\Big)', '\Big(\Big(\big(0+(1\times b)\big)+(0\times e)\Big)+(0\times 0)\Big)', '\Big(\Big(\big(0+(1\times c)\big)+(0\times f)\Big)+(0\times 1)\Big)'].\quad (5)$$

тогда, первое значение ВВХ (вероятность нахождения процесса в поглощающем состоянии S_2 цепи после первого шага $P_{S_2}^{(1)}$) составит:

$$'\Big(\Big(\big(0+(1\times c)\big)+(0\times f)\Big)+(0\times 1)\Big)'. \quad (6)$$

```
t = P_01
for _ in range(elements):
    t = calculate_vector(P, t)
result.append(t)
```

Для того чтобы фрагмент программы реализовал дальнейшие действия, необходимо задать процедуры matrix и coefficients, которые запрашивают конкретные (числовые) значения матриц [P] и [P_01]. Теперь данный фрагмент программы запрашивает процедуру calculate_vector и выполняет ее шесть раз, так как elements =6 (по количеству искомых переходных вероятностей).

Таким образом, рассчитываются в виде уравнений (первое уравнение записано выше), все переходные вероятности 'a', 'b', 'c', 'd', 'e', 'f' записанные в матрицу.

Так как в дальнейшем для определения переходных вероятностей необходимо использовать значения вероятности нахождения системы в поглощающем состоянии - $P_{S_2}^{(1)}$ необходимо использовать выражение (6). Для этого используем оператор цикла for, который будет «перебирать» уравнения списка result и сохранять их в другой список P_02 = [].

Например, из выражения (5) программа сохранит лишь:

$$'\Big(\Big(\big(0+(1\times u)\big)+(0\times k)\Big)+(0\times 1)\Big)'.$$

Реализация данного цикла такова:

```
P_02 = []
for tyt in result:
    ded = tyt[-1]
    P_02.append(ded)
```

Для нахождения искомых переходных вероятностей задаются выходные характеристики (значения требуемых ВВХ, которые определяет полностью или частично заказчик и

доопределяет, по необходимости, исследователь), они вводятся в программу в виде списка, например:

$$VVH = ['0.3', '0.54', '0.702', '0.8076', '0.87588', '0.919944', '0.9483672', '0.96669936', '0.978522768']. \quad (7)$$

Таким образом, следующий фрагмент программы таков:

```
count = len(VVX)
if count != elements+len(P[2]):
    print (f"Ошибка: количество элементов в списке равно {count}, а должно быть {elements+len(P[2])}.")
    exit()

else:
   print ("Количество элементов в списке верное")
```

Данный фрагмент считает количество элементов в списке `VVX` и сравнивает его с количеством значений в матрице (2). Если полученное значение не совпадает с заданным (некорректный ввод пользователя), программа выводит сообщение об ошибке, указывающее фактическое и ожидаемое количество элементов, и завершает работу всей программы. В случае совпадения она выводит сообщение о том, что количество элементов в списке верное.

Так как список (7) задан в строковом виде, то для дальнейшей работы программы для расчета разницы между введенными (заданными) и полученными ВВХ, необходим его перевод в вещественный (математический) вид Float. Это реализуется следующим образом:

```
VVX1 = list(map(float,VVX)).
```
(8)

Далее, продолжим составление уравнений, для этого в полученные уравнения введем значения ВВХ указанные в (7):

```
for k, v in zip(P_02, VVX):
   summed_value = k + '-' + v
   resslt_str.append(summed_value)
print(resslt_str)
```

Данный цикл перебирает поэлементно списки P_02 и VVX и записывает их в список resslt_str = [] в виде нелинейных уравнений, например:

$$'\Big(\big((0+(1\times a))+(0\times f)\big)+(0\times 1)\Big)-0.3'. \quad (9)$$

Для дальнейшей работы программы необходимо представить эти уравнения в вещественном виде, перевод в вещественный вид осуществляет следующий цикл:

```
for i in range(elements):
resslt_strr.append(sp.simplify(resslt_str[i]))
```

Далее составим систему нелинейных уравнений следующим образом:

```
a, c, d, e, f = sp.symbols('a, c, d, e, f ')
eq1 = resslt_strr[0]
eq2 = resslt_strr[1]
eq3 = resslt_strr[2]
eq4 = resslt_strr[3]
eq5 = resslt_strr[4]
eq6 = resslt_strr[5]
eq7 = rr
```

Для расчета системы воспользуемся библиотекой sympy:

```
solution = sp.solve((eq1, eq2, eq3, eq4, eq5, eq6, eq7), (a, c, d, e, f))
```

Для приведения матрицы к квадратному виду, вначале вставим к полученным двум строкам, третью [0 , 0 , 1], реализуется это так:

```
py_list = list(solution[0])
index = py_list.index(0.3)
py_list.insert(index + 1 , 0.3)
py_list.extend([0 , 0 , 1])
solution = tuple([py_list])
fail = list(solution)
```

далее из полученного списка сформируем матрицу:

```
hi1= fail[0][:3]
hi2= fail[0][3:6]
hi3= fail[0][6:9]
PDV = [hi1,hi2,hi3]
```

В результате получена следующая матрица с искомыми значениями переходных вероятностей:

$$\begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.4 & 0.4 \\ 0 & 0 & 1 \end{bmatrix}. \quad (10)$$

Далее проверим - формирует ли требуемые ВВХ найденные значения переходных вероятностей, для этого рассчитаем значения разности между введенными ВВХ и найденными. Для этого рассчитаем ВВХ по найденным значениям переходных вероятностей и вычтем их из заданных значений ВВХ, реализуется это следующим образом:

```
P_03 = [1, 0, 0]
def calculate_vectorin(matrix, coefficients):
   result = []
   for i in range(len(P_01)):
       t_i = 0
       for j in range(len(P_01[0])):
            t_i =  t_i  + (coefficients[j] * matrix[j][i])
       result.append(t_i)
return result
resalt = []
t = P_03
for _ in range(elements):
      t = calculate_vectorin(PDV, t)
      resalt.append(t)
P_04 = []
for tyt in resalt:
     ded = tyt[-1]
     P_04.append(ded)

for:  k, v in zip(P_04, VVX1):
resslt_str1 = []
for k, v in zip(kak, VVX1):
   summed_value = k – v
   resslt_str1.append(summed_value)
```

В результате получено:

$$[0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0] \tag{11}$$

- значения полностью совпали, следовательно, переходные вероятности программой определены корректно.

Таким образом, создана программа реализующая решение обратной задачи по нахождению переходных вероятностей поглощающей конечной марковской цепи по динамики вероятности поглощающего состояния на базе языка программирования Python в среде Visual Studio. Программа представляет собой инструмент решения обратной задачи в терминах марковских процессов, она может быть легко адаптирована под конкретную конечную марковскую цепь [2,5].

Литература

1. Постановка задачи исследования процесса повышения оперативности доведения информации на сеансе связи в перспективном пакетном декаметровом радиоканале АСУ общего назначения Попов М. Ю., Девятияров В. Б., Торбенко Я. К., Киреев Д. А. В сборнике: Радиоэлектронные устройства и системы для инфокоммуникационных технологий" ("РЭУС-ИТ 2024"). Доклады Всероссийской конференции, посвященной Дню радио. Москва, 2024. С. 129-133.

2. Математическая модель процесса доведения сообщения по радиоканалу управления специальной АСУ в условиях энергетического и информационного воздействия на базе конечной марковской цепи с доходами Цимбал В. А., Попов М. Ю., Илющенко В. В., Калганов Е. В., Деркач А. М. В сборнике: Цифровая обработка сигналов и её применение (DSPA-2022). 24-я Международная конференция. Москва, 2022. С. 70-75.

3. Численный анализ характеристик неоднородных конечных марковских цепей при разных шагах переходов Цимбал В. А., Попов М. Ю., Якимова И. А. В сборнике: Материалы конференции "Информационные технологии в управлении" (ИТУ-201 4). 2014. С. 738-742.

4. Численное определение временных и вероятностно-временных характеристик поглощающих конечных марковских цепей при конечном числе шагов процесса Бородин С. Г., Деркач А. М., Дробышев М. Ю., Попов М. Ю., Цимбал В. А. REDS: Телекоммуникационные устройства и системы. 2014. Т. 4. № 2. С. 115-119.

5. Математическое моделирование процесса доведения сообщения в радиосети без обратной связи с повторениями и накоплением информации Цимбал В. А., Попов М. Ю., Дробышев М. Ю. Информационные технологии в проектировании и производстве. 2010. № 3. С. 78-83.

THE DECISION OF THE RETURN PROBLEM ON THE BASIS OF THE ABSORBING FINAL MARKOVSKY CHAIN BY MEANS OF PROGRAMMING LANGUAGE PYTHON IN THE ENVIRONMENT OF VISUAL STUDIO

Dr.Sci.Tech. Popov M.Yu.[1], Cand.Tech.Sci. Devyatiyarov V.B.[2],
Cand.Tech.Sci. Slobodskov O.E.[3], Ponomarev D.V.[1]

[1]Branch of military academy of Rocket strategic forces of a name of Peter the Great (c. Serpukhov, Moscow region).
[2]Joint-stock company corporation «Moscow institute of heat engineering».
[3]The Russian institute of a powerful radio structure, St.-Petersburg.

In article the approach on a finding of transitive probabilities of a matrix of transitive probabilities of absorbing final chain Markova on preset values of probabilities of an absorbing condition is presented. The corresponding settlement program is created. The program is realised in programming language Python in the environment of Visual Studio. The program can be used at the decision of a wide range of return problems by definition of entrance parametres of processes formalized by absorbing final chains Markova on days off.

МАТЕМАТИЧЕСКИЕ МОДЕЛИ ПРОЦЕССА ДОВЕДЕНИЯ СООБЩЕНИЯ В НИЗКОЧАСТОТНОМ ДУПЛЕКСНОМ РАДИОКАНАЛЕ АСУ ОБЩЕГО НАЗНАЧЕНИЯ НА БАЗЕ НАУЧНО-МЕТОДИЧЕСКОГО АППАРАТА КОНЕЧНЫХ ЦЕПЕЙ МАРКОВА

д.т.н., доц. Попов М.Ю., асп. Торбенко Я.К., асп. Чернухин А.В.

Филиал военной академии РВСН имени Петра Великого, г. Серпухов

В статье представлены графы состояний и переходов поглощающей и регулярной конечных цепей Маркова, описывающие процессы доведения сообщения в дуплексном низкочастотном радиоканале управления автоматизированной системы управления общего назначения. Приведен физический смысл состояний цепей, переходных вероятностей и шагов переходов. Регулярная и поглощающая марковские цепи процесса являются базой для построения соответствующих математических моделей и служат инструментом получения вероятностно-временных характеристик.

Известна система централизованного оповещения, построенная на базе низкочастотного (НЧ) радиоканала управления, функционирующего в условиях помех. Радиоканал является симплексным и служит для доставки сообщений различной длины l_c. Передача сообщения от объекта управления (ОУ) осуществляется НЧ передающим радиоцентром способом простого повторения на конкретной (выбранной) рабочей частоте (РЧ). Причем, доведение сообщения может осуществляться поочередно на нескольких РЧ. Например, сначала на одной РЧ (передается первая часть повторов), а после - на второй (вторая часть повторов). Повторы сообщения при доведении имеют одинаковую длину, которая устанавливается в начале передачи и зависит от типа доставляемой информации.

Симплексный НЧ радиоканал способен функционировать в режиме избирательной доставки сообщения (соединение «точка-точка») и обеспечивать требуемые вероятностно-временные характеристики (ВВХ) за конечное время работы НЧ передающего (ПРД) средства, не зависящее от вероятности доставки сообщения [5, 6].

Вместе с тем существуют требования по снижению времени работы НЧ ПРД средств (экономия ограниченного гарантийного времени работы НЧ передатчиков) при обеспечении требуемых ВВХ доставки сообщения. Обеспечить данные требования возможно путем организации дуплексного режима функционирования рассматриваемого радиоканала. Для этого необходимо создать соответствующие математические модели рассматриваемого процесса, для чего введем следующие определения:

Определение 1. Сообщение 1 типа – сообщение, информация которого полностью «помещается» в информационную часть одного повтора.

Определение 2. Сообщение 2 типа – сообщение, информация которого «помещается» в информационные части двух повторов (доставляется двумя повторами).

Определение 3. Квитанция 1 типа – квитанция, информация которой полностью «помещается» в один блок информационной части.

Определение 4. Квитанция 2 типа – квитанция, информация которой «помещается» в два блока информационной части.

На рисунках 1 и 2 приведены графы состояний и переходов процесса доведения сообщения 1 типа по НЧ дуплексному радиоканалу в виде поглощающей и регулярной конечных марковских цепей (ПКМЦ, РКМЦ).

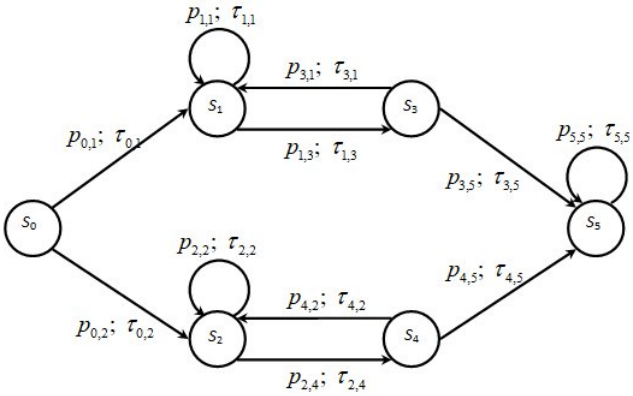


Рисунок 1 – ПКМЦ процесса доведения сообщения 1 типа

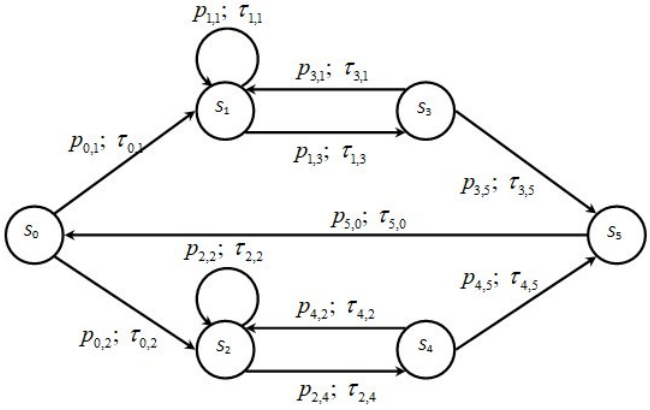


Рисунок 2 – РКМЦ процесса доведения сообщения 1 типа

Физическая сущность состояний цепи:

S_0 – наличие сообщения у отправителя для доставки получателю;

S_1 – передача отправителем сообщения длиной l_c получателю (V_{f1});

S_2 – передача отправителем сообщения l_c получателю (V_{f2});

S_3 – доставка получателю сообщения l_c от отправителя (V_{f1}). Передача получателем квитанции l_{k_1} отправителю (V_{f2}).

S_4 – доставка получателю сообщения l_c от отправителя (V_{f2}). Передача получателем квитанции l_{k_1} отправителю (V_{f1}).

S_5 – доставка квитанции от получателя отправителю – поглощающее состояние.

Физическая сущность переходных вероятностей шагов перехода цепей:

$p_{0,1};\tau_{0,1}$ – организация передачи сообщения отправителем на f_1 получателю;

$p_{0,2};\tau_{0,2}$ – передача сообщения отправителем на f_2 получателю;

$p_{1,3};\tau_{1,3}$ – доставка сообщения l_c получателю от отправителя (V_{f1});

$p_{1,1};\tau_{1,1}$ – недоставка сообщения l_c получателю от отправителя (V_{f1});

$p_{2,4};\tau_{2,4}$ – доставка сообщения l_c получателю от отправителя (V_{f2});

$p_{2,2};\tau_{2,2}$ – недоставка сообщения l_c получателю от отправителя (V_{f2});

$p_{3,5};\tau_{3,5}$ – доставка квитанции l_{k_1} отправителю от получателя (V_{f2});

$p_{3,1};\tau_{3,1}$ – недоставка квитанции l_{k_1} отправителю от получателя (V_{f2});

$p_{4,5};\tau_{4,5}$ – доставка квитанции l_{k_1} отправителю от получателя (V_{f1});

$p_{4,2};\tau_{4,2}$ – недоставка квитанции l_{k_1} отправителю от получателя (V_{f1});

$p_{5,5};\tau_{5,5}$ – поглощающее состояние;

$p_{5,0};\tau_{5,0}$ – возврат к состоянию S_0.

МПВ имеет вид:

$$P_{[6,6]} = \begin{Vmatrix} 0 & p_{0,1} & p_{0,2} & 0 & 0 & 0 \\ 0 & p_{1,1} & 0 & p_{1,3} & 0 & 0 \\ 0 & 0 & p_{2,2} & 0 & p_{2,4} & 0 \\ 0 & p_{3,1} & 0 & 0 & 0 & p_{3,5} \\ 0 & 0 & p_{4,2} & 0 & 0 & p_{4,5} \\ 0 & 0 & 0 & 0 & 0 & p_{5,5} \end{Vmatrix}. \quad (1)$$

МШП имеет вид:

$$T_{[6,6]} = \begin{Vmatrix} 0 & t_{0,1} & t_{0,2} & 0 & 0 & 0 \\ 0 & t_{1,1} & 0 & t_{1,3} & 0 & 0 \\ 0 & 0 & t_{2,2} & 0 & t_{2,4} & 0 \\ 0 & t_{3,1} & 0 & 0 & 0 & t_{3,5} \\ 0 & 0 & t_{4,2} & 0 & 0 & t_{4,5} \\ 0 & 0 & 0 & 0 & 0 & t_{5,5} \end{Vmatrix}. \quad (2)$$

ПКМЦ служит инструментом определения ВВХ и временных характеристик процесса посредством соответствующего численного анализа. Под ВВХ можно понимать динамику вероятности доведения сообщения от времени, а под вероятностными характеристиками (ВХ) – среднее время доведения и среднее квадратическое отклонение (СКО) времени.

РКМЦ необходима для оценки временных характеристик работы соответствующих передающих средств.

На рисунке 3 представлены графы состояний и переходов ПКМЦ и РКМЦ, описывающие процесс доведения сообщений 2 типа.

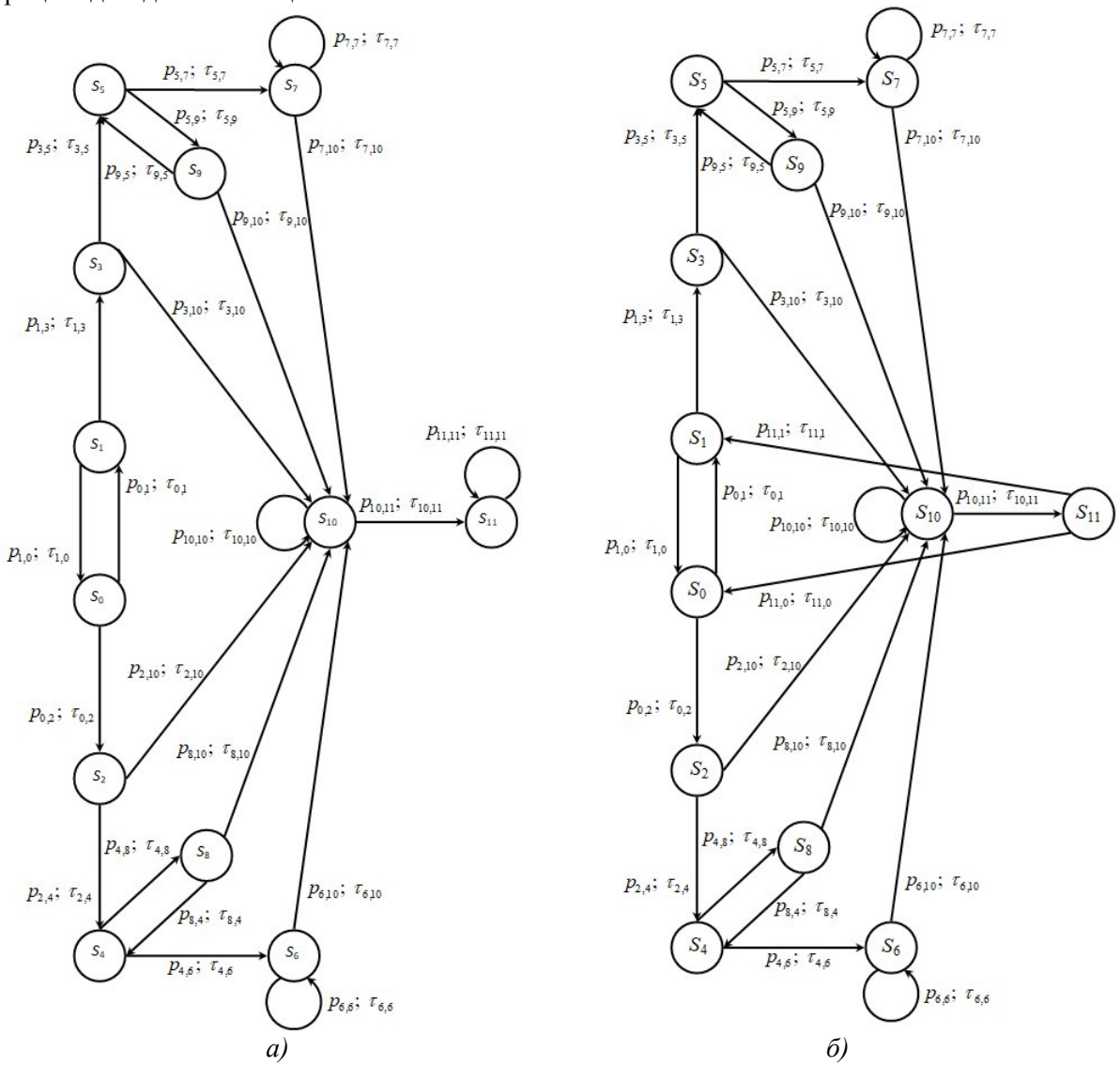


Рисунок 3 – *а)* ПКМЦ процесса доведения сообщения 2 типа; *б)* РКМЦ процесса доведения сообщения 2 типа

Физическая сущность состояний цепи:

S_0 – передача 1 повтора сообщения длиной l_{c_1} отправителем (V_{f1}) получателю;

S_1 – передача 2 повтора l_{c_2} отправителем (V_{f1}) получателю;

S_2 – доставка получателю 1 повтора l_{c_1} (V_{f1}) от отправителя. Передача 2 повтора l_{c_2} отправителем (V_{f1}) получателю;

S_3 – доставка получателю 2 повтора l_{c_2} (V_{f1}) от отправителя. Передача 1 повтора l_{c_1} отправителем (V_{f1}) получателю;

S_4 – доставка получателю 1 повтора l_{c_1} и недоставка 2 повтора l_{c_2} (V_{f1}) от отправителя. Передача получателем квит. 1 типа l_{k_1} отправителю (V_{f2});

S_5 – доставка получателю 2 повтора l_{c_2} и недоставка 1 повтора l_{c_1} (V_{f1}) от отправителя. Передача получателем квит. 1 типа l_{k_1} отправителю (V_{f2});

S_6 – доставка получателю 1 повтора l_{c_1} и недоставка 2 повтора l_{c_2} (V_{f1}) от отправителя. Доставка отправителю квитанции 1 типа l_{k_1} получателем (V_{f2}). Передача 2 повтора l_{c_2} отправителем (V_{f1}) получателю;

S_7 – доставка получателю 2 повтора l_{c_2} и недоставка 1 повтора l_{c_1} (V_{f1}) от отправителя. Доставка отправителю квитанции 1 типа l_{k_1} получателем (V_{f2}). Передача 1 повтора l_{c_1} отправителем (V_{f1}) получателю;

S_8 – недоставка отправителю квитанции 1 типа l_{k_1} получателем (V_{f2}). Продолжение последовательной передачи отправителем 2 повтора l_{c_2} и 1 повтора l_{c_1} (V_{f1}) получателю;

S_9 – недоставка отправителю квитанции 1 типа l_{k_1} получателем (V_{f2}). Продолжение последовательной передачи отправителем 1 повтора l_{c_1} и 2 повтора l_{c_2} (V_{f1}) получателю;

S_{10} – передача получателем квитанции 2 типа l_{k_2} отправителю (V_{f2});

S_{11} – доставка от получателя квитанции 2 типа l_{k_2} отправителю (V_{f2}) – поглощающее состояние.

Физическая сущность переходных вероятностей и шагов переходов:

$p_{0,1};\tau_{0,1}$ – недоставка 1 повтора l_{c_1} получателю от отправителя (V_{f1});

$p_{0,2};\tau_{0,2}$ – доставка 1 повтора l_{c_1} получателю от отправителя (V_{f1});

$p_{1,0};\tau_{1,0}$ – недоставка 2 повтора l_{c_2} получателю от отправителя (V_{f1});

$p_{1,3};\tau_{1,3}$ – доставка 2 повтора l_{c_2} получателю от отправителя (V_{f1});

$p_{2,4};\tau_{2,4}$ – недоставка 2 повтора l_{c_2} получателю от отправителя (V_{f1});

$p_{2,10};\tau_{2,10}$ – доставка 2 повтора l_{c_2} получателю от отправителя (V_{f1});

$p_{3,5};\tau_{3,5}$ – недоставка 1 повтора l_{c_1} получателю от отправителя (V_{f1});

$p_{3,10};\tau_{3,10}$ – доставка 1 повтора l_{c_1} получателю от отправителя (V_{f1});

$p_{4,8};\tau_{4,8}$ – недоставка квит. 1 типа l_{k_1} отправителю от получателя (V_{f2});

$p_{4,6};\tau_{4,6}$ – доставка квит. 1 типа l_{k_1} отправителю от получателя (V_{f2});

$p_{5,9};\tau_{5,9}$ – недоставка квит. 1 типа l_{k_1} отправителю от получателя (V_{f2});

$p_{5,7};\tau_{5,7}$ – доставка квит. 1 типа l_{k_1} отправителю от получателя (V_{f2});

$p_{6,10};\tau_{6,10}$ – доставка 2 повтора l_{c_2} получателю от отправителя (V_{f1});

$p_{3,6};\tau_{3,6}$ – недоставка 2 повтора l_{c_2} получателю от отправителя (V_{f1});

$p_{7,10};\tau_{7,10}$ – доставка 1 повтора l_{c_1} получателю от отправителя (V_{f1});

$p_{7,7};\tau_{7,7}$ – недоставка 1 повтора l_{c_1} получателю от отправителя (V_{f1});

$p_{8,10};\tau_{8,10}$ – доставка 2 повтора l_{c_2} получателю от отправителя (V_{f1}) в результате последовательной передачи отправителем 1 повтора l_{c_1} и 2 повтора l_{c_2} (V_{f1});

$p_{8,4};\tau_{8,4}$ – недоставка 2 повтора l_{c_2} получателю от отправителя (V_{f1}) в результате последовательной передачи отправителем 1 повтора l_{c_1} и 2 повтора l_{c_2} (V_{f1});

$p_{9,10};\tau_{9,10}$ – доставка 1 повтора l_{c_1} получателю от отправителя (V_{f1}) в результате последовательной передачи отправителем 1 повтора l_{c_1} и 2 повтора l_{c_2} (V_{f1});

$p_{9,5};\tau_{9,5}$ – недоставка 1 повтора l_{c_1} получателю от отправителя (V_{f1}) в результате последовательной передачи отправителем 1 повтора l_{c_1} и 2 повтора l_{c_2} (V_{f1});

$p_{10,11};\tau_{10,11}$ – доставка квит. 2 типа l_{k_2} отправителю от получателя (V_{f2});

$p_{10,10};\tau_{10,10}$ – недоставка квитанции 2 типа l_{k_2} отправителю от получателя (V_{f2});

$p_{11,11};\tau_{11,11}$ – поглощающее состояние;

$p_{11,0};\tau_{11,0}$ – переход к передаче первого повтора;
$p_{11,1};\tau_{11,1}$ – переход к передаче второго повтора.
МПВ имеет вид:

$$P_{[12,12]} = \begin{Vmatrix} 0 & p_{0,1} & p_{0,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_{1,0} & 0 & 0 & p_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{2,4} & 0 & 0 & 0 & 0 & 0 & p_{2,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{3,5} & 0 & 0 & 0 & 0 & p_{3,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & p_{4,6} & 0 & p_{4,8} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{5,7} & 0 & p_{5,9} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & p_{6,6} & 0 & 0 & 0 & p_{6,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{7,7} & 0 & 0 & p_{7,10} & 0 \\ 0 & 0 & 0 & 0 & p_{8,4} & 0 & 0 & 0 & 0 & 0 & p_{8,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{9,5} & 0 & 0 & 0 & 0 & p_{9,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{10,10} & p_{10,11} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{11,11} \end{Vmatrix}. \quad (3)$$

МШП имеет вид:

$$T_{[12,12]} = \begin{Vmatrix} 0 & t_{0,1} & t_{0,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ t_{1,0} & 0 & 0 & t_{1,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & t_{2,4} & 0 & 0 & 0 & 0 & 0 & t_{0,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & t_{3,5} & 0 & 0 & 0 & 0 & t_{3,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & t_{4,6} & 0 & t_{4,8} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_{5,7} & 0 & t_{5,9} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & t_{6,6} & 0 & 0 & 0 & t_{6,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_{7,7} & 0 & 0 & t_{7,10} & 0 \\ 0 & 0 & 0 & 0 & t_{8,4} & 0 & 0 & 0 & 0 & 0 & t_{8,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & t_{9,5} & 0 & 0 & 0 & 0 & t_{9,10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_{10,10} & t_{10,11} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_{11,11} \end{Vmatrix}. \quad (4)$$

После определения ПВ и ШП вся информация о процессе будет содержаться в данных матрицах. С помощью их численного анализа возможно получить характеристики, интересующие исследователя [1-4].

Таким образом, на базе научно-методического аппарата конечных цепей Маркова созданы поглощающие и регулярные марковские цепи, описывающие процесс доведения сообщения в дуплексном НЧ радиоканале управления АСУ общего назначения.

Литература

1. Аналитические выражения и программные средства определения числовых характеристик конечных цепей Маркова. Попов М. Ю., Девятияров В. Б., Слободсков О. Е., Харитонов И. А. Проектирование и технология электронных средств. 2025. № 1. С. 48-57.

2. Определение показателей динамики процесса, формализированного поглощающей или регулярной конечной полумарковской цепью в зависимости от времени на базе языка программирования С++ в среде Visual Studio. Попов М. Ю., Калганов Е. В., Асанов А. С., Харитонов И. А. В сборнике: Радиоэлектронные устройства и системы для инфокоммуникационных технологий" ("РЭУС-ИТ 2024"). Доклады Всероссийской конференции, посвященной Дню радио. Москва, 2024. С. 114-118.

3. Программа расчета численных характеристик регулярной конечной полумарковской цепи на базе языка программирования С++ в среде Visual Studio. Попов М. Ю., Илюшенко В. В., Калинин Д. С., Торбенко Я. К., Харитонов И. А. В сборнике: Цифровая обработка сигналов и ее применение DSPA - 2024. Доклады XXVI Международной конференции. Москва, 2024. С. 23-28.

4. Программа расчета темповых характеристик регулярной конечной марковской цепи на базе языка программирования С++ в среде Visual Studio. Попов М. Ю., Ржаных А. В., Киреев Д.

А., Харитонов И. А. В сборнике: Цифровая обработка сигналов и ее применение DSPA - 2024. Доклады XXVI Международной конференции. Москва, 2024. С. 276-280.

5. Численный анализ характеристик неоднородных конечных марковских цепей при разных шагах переходов. Цимбал В. А., Попов М. Ю., Якимова И. А. В сборнике: Материалы конференции "Информационные технологии в управлении" (ИТУ-2014). 2014. С. 738-742.

6. Математическая модель процесса доведения сообщения по радиоканалу управления специальной АСУ как инструмент моделирования информационного конфликта. Цимбал В. А., Попов М. Ю., Илющенко В. В., Калганов Е. В., Киреев Д. А. В сборнике: цифровая обработка сигналов и её применение (DSPA-2022). 24-я Международная конференция. Москва, 2022. С. 26-31.

MATHEMATICAL MODELS OF THE PROCESS OF MESSAGE DELIVERY IN A LOW-FREQUENCY DUPLEX RADIO CHANNEL OF A GENERAL-PURPOSE ACS BASED ON THE SCIENTIFIC AND METHODOLOGICAL APPARATUS OF FINITE MARKOV CHAINS

dr.sci.tech. Popov M.Yu., grad.st. Torbenko Ya.K., grad.st. Chernukhin A.V.

Branch of military academy of Rocket strategic forces of a name of Peter the Great
(c. Serpukhov, Moscow region).

The article presents graphs of states and transitions of absorbing and regular finite Markov circuits describing the processes of delivering a message in a duplex low-frequency radio control channel of a general-purpose automated control system. The physical meaning of circuit states, transition probabilities, and transition steps is given. The regular and absorbing Markov chains of the process are the basis for constructing appropriate mathematical models and serve as a tool for obtaining probabilistic and temporal characteristics.

МЕТОДИКА ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ НИЗКОЧАСТОТНОГО ДУПЛЕКСНОГО РАДИОКАНАЛА АСУ ОБЩЕГО НАЗНАЧЕНИЯ, ОБЕСПЕЧИВАЮЩИХ ТРЕБУЕМУЮ ВЕРОЯТНОСТЬ ДОВЕДЕНИЯ СООБЩЕНИЯ

асп. Торбенко Я.К.

Филиал военной академии РВСН имени Петра Великого, г. Серпухов

В статье представлена формальная постановка задачи по определению параметров дуплексного низкочастотного радиоканала управления автоматизированной системы управления общего назначения, обеспечивающих требуемую вероятность доставки сообщения. В статье рассматриваются скоростные параметры дуплексного радиоканала, под которыми понимаются соответствующие скорости передачи сообщений на различных рабочих частотах. Показаны взаимосвязь этапов исследования, этапы методики, необходимые для решения поставленной задачи, а также ее блок-схема.

В настоящее время существуют требования по снижению времени работы низкочастотных (НЧ) передающих средств симплексного радиоканала управления АСУ общего назначения при обеспечении требуемой вероятности избирательной доставки сообщения в условиях помех. Обеспечить выполнение указанных требований возможно посредством организации дуплексного НЧ радиоканала и обоснования его параметров. Это возможно сделать с помощью соответствующей методики.

Назначение методики - определение скоростей передачи повторов сообщений, доставляемых объектом управления (ОУ) - отправителем на соответствующих рабочих частотах (РЧ) в дуплексном радиоканале управления АСУ общего назначения. Скоростные параметры должны обеспечить минимизацию суммарного времени работы передающих НЧ радиосредств в дуплексном режиме, при этом обеспечить требуемую вероятность доведения сообщения в условиях помех.

Соответствующая формальная постановка задачи имеет вид:

$$\begin{gathered}\left\{V_{C_i};V_{K_i}\right\}:\\ T_{раб.ПРД}^{симпл.}-\sum T_{раб.ПРД}^{дупл.}(V_{C_i},V_{K_i})\to \max\\ где\ \ T_{раб.ПРД}^{симпл.}=const\\ \Rightarrow \sum T_{раб.ПРД}^{дупл.}(V_{C_i},V_{K_i})\to \min\\ при\ \ P_{дост.}^{дупл.}(V_{C_i},V_{K_i})=P_{дост.}^{симпл.}(V_{C_i});\\ P_{дов.}^{дупл.}(V_{C_i},V_{K_i})\geq P_{дов.}^{дупл.треб}(V_{C_i},V_{K_i}).\end{gathered} \qquad (1)$$

При решении данной задачи необходимо учитывать следующее теоретическое противоречие, которое схематично показано на рисунке 1, где: V – скорость передачи сообщения [бит/с], t – время доведения сообщения [с], h_0^2 - отношение сигнал/шум [дБ], p_0 – вероятность ошибки на входе демодулятора, $P_{дов}$ – вероятность доведения сообщения.

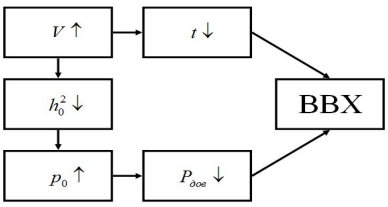


Рисунок 1 – Теоретическое противоречие

На рисунке 2 представлена логическая взаимосвязь этапов исследования.



Рисунок 2 – Взаимосвязь этапов исследования

Этапы методики:

1. Формирование исходных данных, с учетом существующих ограничений, допущений и требований.

1.1 Модельных параметров.

1.2 Параметров радиоканала, сообщений и квитанций.

2 Формирование переходных вероятностей и шагов перехода поглощающей конечной марковской цепи описывающей процесс.

3 Формирование матрицы переходных вероятностей (МПВ).

3.1 Проверка стохастичности МПВ.

4. Формирование матрицы шагов перехода.

5. Определение вероятностно-временных характеристик (ВВХ) процесса.

5.1 Определение ВВХ в динамике шагов процесса.

5.1.1 Формирование вектора начальных состояний.

5.1.2 Формирование уравнений Колмогорова-Чепмена (УКЧ) в матричной форме.

5.2 Определение ВВХ в динамике времени процесса.

5.2.1 Формирование вектора-столбца частных средних шагов перехода.

5.2.2 Формирование единой оси времени.

5.2.3 Формирование функций ВВХ.

5.2.4 Расчет и построение графиков ВВХ.

5.3 Нахождение максимального значения вероятности доведения сообщения и соответствующего ему времени.

6. Проверка достаточности заданного исследователем количества шагов процесса моделирования.

7. Определение вероятностных характеристик (ВХ) процесса.

7.1 Определение средних времен достижения состояний процесса численным, оценочным способом.

7.1.1 Расчет среднего времени доведения сообщения.

7.1.2 Расчет среднего квадратического отклонения (СКО) времени доведения.

8. Определение степени загруженности состояний процесса на базе РКМЦ.

8.1 Формирование регулярной конечной марковской цепи (РКМЦ).

8.1.1 Формирование регулярной переходной матрицы.

8.1.2 Формирование вектора-столбца частных средних шагов перехода РКМЦ.

8.2 Определение предельного вектора.

8.3 Проверка стохастичности предельного вектора.

8.4 Определение доли времени нахождения процесса в состояниях РКМЦ - оценка степени «загруженности» состояний процесса.

8.4.1 Расчет времени нахождения процесса в состояниях цепи.

8.4.2 Проверка адекватности формирования вектора времен нахождения процесса в состояниях цепи.

8.4.3 Расчет времени (и процента) работы передатчиков ОУ – отправителя.

8.4.4 Расчет времени (и процента) работы передатчиков ОУ – получателя.

8.4.5 Расчет общего времени (и процента) работы передатчиков ОУ.

9. Определение эффекта.

9.1 Расчет времени работы передающего средства ОУ - отправителя в одностороннем радиоканале.

9.2 Расчет выигрыша во времени (и в %) работы передающего средства ОУ - отправителя при доведении сообщения по дуплексному радиоканалу.

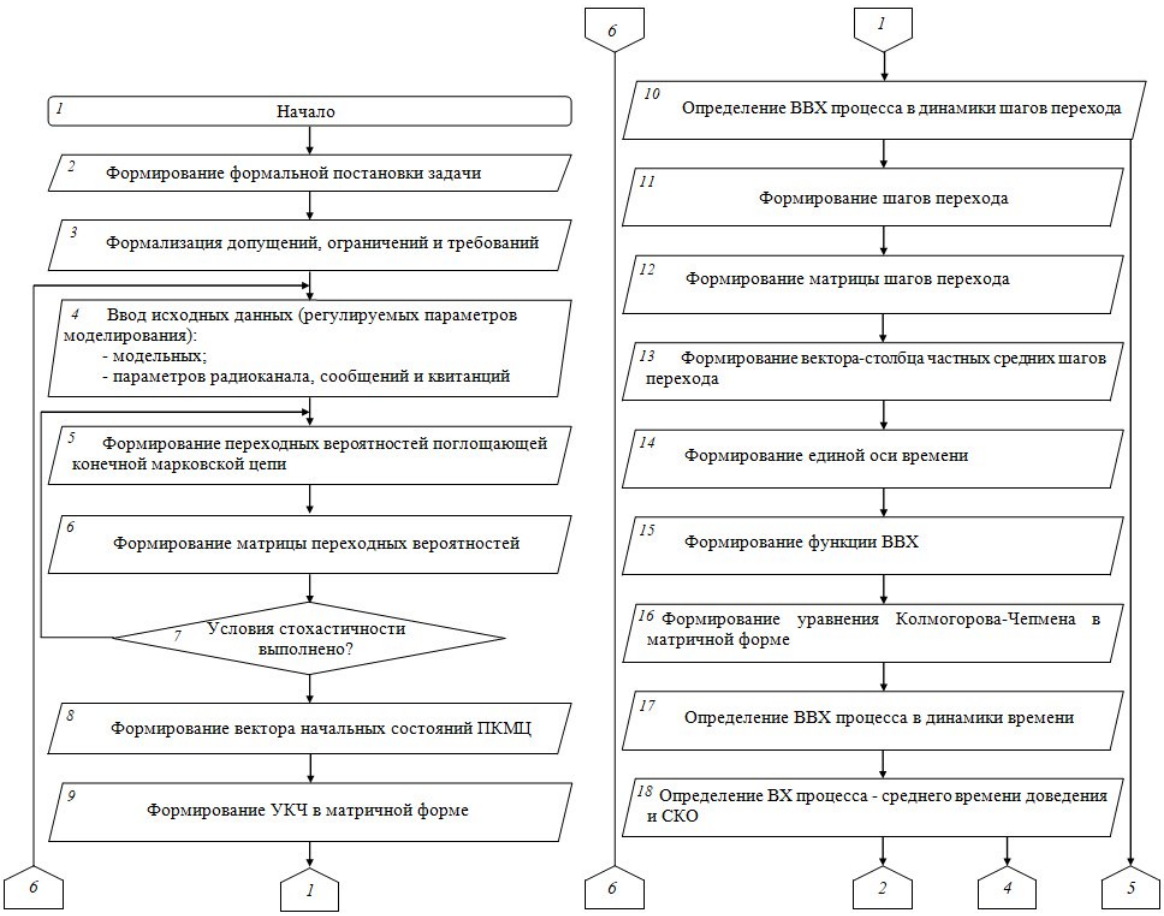
9.3 Расчет выигрыша во времени (и в %) работы передающих средств ОУ (отправителя и получателя) при доведении сообщения по дуплексному радиоканалу.

10. Анализ полученных данных и интерпретация их в физической сущности процесса.

11. Выявления закономерностей поведения физического процесса.

12. Формирование научно-обоснованных рекомендаций.

Указанные этапы взаимоувязаны в блок-схему, изображенную на рисунке 3.



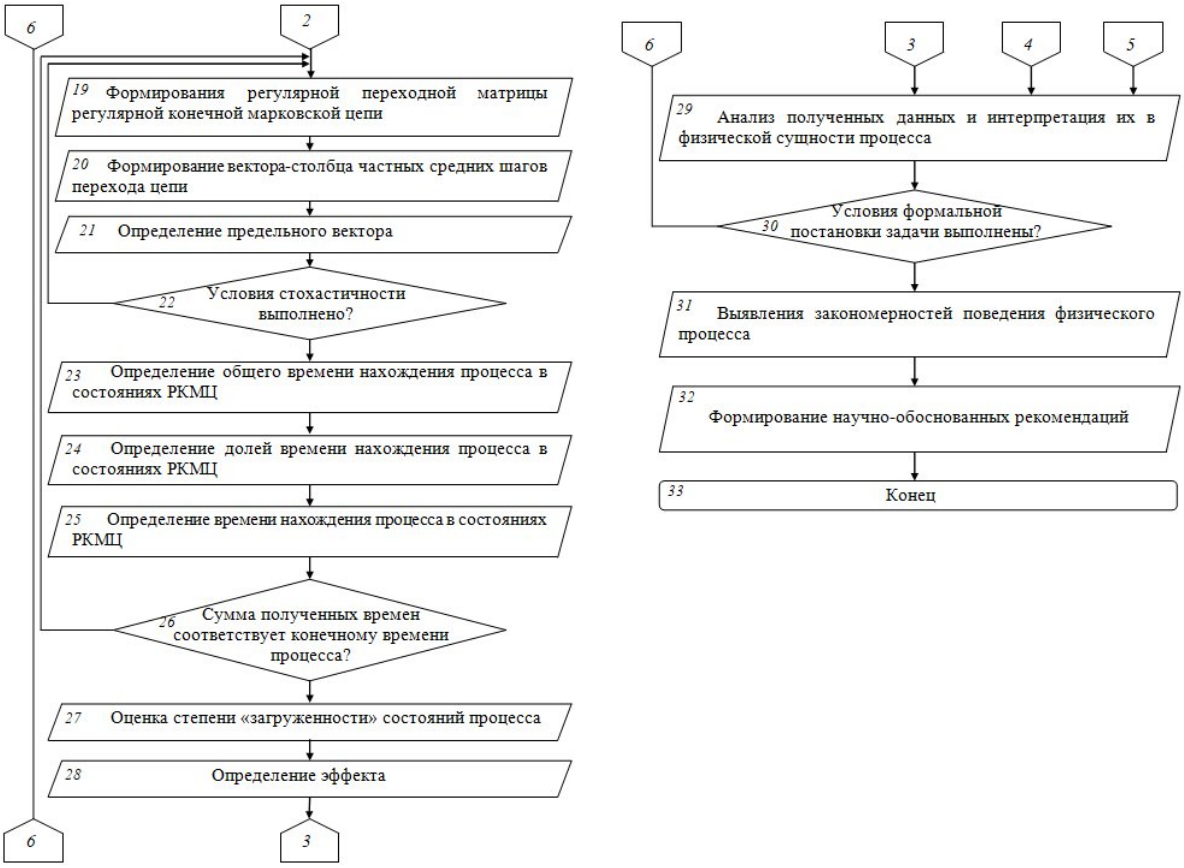


Рисунок 3 – Блок-схема методики

Методика является инструментом получения характеристик информационного обмена и временных характеристик работы низкочастотных передающих средств объектов управления [1-4].

Практическая значимость методики заключается в возможности формировать обоснованные предложения по совершенствованию НЧ приемо-передающей аппаратуры (ППА) и аппаратуры передачи данных (АПД) комплекса технических средств (КТС) АСУ общего назначения с целью снижения времени работы НЧ ПРД средств при обеспечении требуемых ВВХ доведения сообщения;

Применение методики позволит сформировать рекомендации (например предприятиям промышленности) по совершенствованию математического (алгоритмического и программного) обеспечения НЧ ППА и АПД КТС АСУ общего назначения, а также органам управления по порядку использования предлагаемого варианта дуплексного радиоканала в режиме избирательной доставки сообщений большого объема для межобластного взаимодействия.

Литература

1. Методика оценки загруженности состояний процесса, формализованного регулярной полумарковской цепью. Попов М. Ю., Девятияров В. Б., Слободсков О. Е., Калганов Е. В., Шаблюк С. М. В сборнике: Цифровая обработка сигналов и ее применение DSPA - 2024. Доклады XXVI Международной конференции. Москва, 2024. С. 18-22.

2. Численное определение временных и вероятностно-временных характеристик поглощающих конечных марковских цепей при конечном числе шагов процесса. Бородин С. Г., Деркач А. М., Дробышев М. Ю., Попов М. Ю., Цимбал В. А. REDS: Телекоммуникационные устройства и системы. 2014. Т. 4. № 2. С. 115-119.

3. Программа расчета численных характеристик регулярной конечной полумарковской цепи на базе языка программирования С++ в среде Visual Studio. Попов М. Ю., Илюшенко В. В.,

Калинин Д. С., Торбенко Я. К., Харитонов И. А. В сборнике: Цифровая обработка сигналов и ее применение DSPA - 2024. Доклады XXVI Международной конференции. Москва, 2024. С. 23-28.

4. Программа расчета темповых характеристик регулярной конечной марковской цепи на базе языка программирования C++ в среде Visual Studio. Попов М. Ю., Ржаных А. В., Киреев Д. А., Харитонов И. А. В сборнике: Цифровая обработка сигналов и ее применение DSPA - 2024. Доклады XXVI Международной конференции. Москва, 2024. С. 276-280.

METHOD FOR DETERMINING THE PARAMETERS OF A LOW-FREQUENCY DUPLEX RADIO CHANNEL OF A GENERAL-PURPOSE AUTOMATED CONTROL SYSTEM, PROVIDING THE REQUIRED PROBABILITY OF DELIVERING A MESSAGE

grad.st. Torbenko Ya.K.

Branch of military academy of Rocket strategic forces of a name of Peter the Great
(c. Serpukhov, Moscow region)

The article presents a formal formulation of the problem of determining the parameters of a duplex low-frequency radio control channel of a general-purpose automated control system that ensure the required probability of message delivery. The article discusses the speed parameters of a duplex radio channel, which are understood as the corresponding message transmission rates at various operating frequencies. The interrelation of the research stages, the stages of the methodology necessary to solve the task, as well as its flowchart are shown.

ОПТИМАЛЬНОЕ НАЧАЛЬНОЕ ПРИБЛИЖЕНИЕ ДЛЯ ЧИСЛЕННОГО ПОИСКА ПАРАМЕТРОВ ДВУМЕРНОГО ГАУССОВА КУПОЛА

асп. Никитин О.В.[1], преп. Овчинников П.Е.[1]

[1] Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского

Рассматривается задача численного поиска положения среднего двумерного Гауссова купола по измеренным значениям функции в наборе точек на некоторой сетке. Решение по методу наименьших квадратов находится с помощью итеративной оптимизации. Для ускорения процедуры поиска предложена методика вычисления оптимального начального приближения положения среднего, основанная на линеаризации системы уравнений. Применение нового начального приближения при низкой плотности дискретизации уменьшает время работы алгоритма на 33% в сравнении с начальным приближением наибольшего значения.

Постановка задачи

Задача поиска параметров Гауссовых куполов по измеренным значениям функции актуальна в различных областях науки. Она встречается в статистике – при поиске распределения измеренной случайной величины по измеренным значениям плотности вероятности. Также она появляется в радиолокации и компьютерном зрении – при выделении признаков на растровом изображении, например, для определения положения небесного тела по его световому пятну.

Математически задача формулируется как система m нелинейных уравнений в n измерениях:

$$f_k = f(\vec{x}_k) = A\exp(-(\vec{x}_k - \vec{\mu})^T\, C\, (\vec{x}_k - \vec{\mu}))), \quad (1)$$

где $k \in [1; m]$ – индекс значения, $\vec{x}_k$ – вектор координат точки, f_k – значение распределения в этой точке, $\vec{\mu}$ – вектор n координат среднего значения распределения, C – положительно определённая матрица $n \times n$ коэффициентов – параметров масштаба распределения.

Численное решение этой системы находится методом наименьших квадратов; классический способ его поиска – итеративная минимизация функционала отклонения, например, методом градиентного спуска по параметрам распределения. Однако поскольку аппроксимационная функция нелинейная, то скорость сходимости и возможность попадания в глобальный минимум функционала сильно зависят от выбора начального приближения. Чем ближе к истинным значениям параметров будет выбрано начальное приближение $\vec{\mu}$ и C, тем лучше будет сходиться алгоритм оптимизации.

Для поиска решения по измеренным значениям $f_k = f(\vec{x}_k)$ также нужно сделать допущение малости шума, иначе поиск аппроксимации не будет сходиться к истинным значениям параметров.

Начальное приближение по максимуму значения

Поскольку квадратичная форма $(\vec{x}_k - \vec{\mu})^T\, C\, (\vec{x}_k - \vec{\mu})$ в показателе экспоненты положительно определённая, распределение достигает своего максимума при наименьшем отклонении координат $\vec{x}_k$ от среднего $\vec{\mu}$. Следовательно, координаты среднего $\vec{\mu}$ должны находиться рядом с координатами точки $\vec{x}_m$ с максимальным известным значением f_m. Таким образом, если среди всех измеренных значений чётко выделяется один максимум, то его следует взять в качестве начального приближения $\vec{\mu}_0 = \vec{x}_m$, и взять его значение в качестве начального приближения $A_0 = f_m$. Если же присутствует несколько значений, близких к максимуму, то в качестве начального приближения $\vec{\mu}_0$ следует взять результат усреднения соответствующих координат.

Преимущество такого приближения – оно быстро вычисляется. Но недостаток в том, что его нельзя применять в случае удалённости истинного положения среднего от всех точек с измеренными значениями функции. Подобная ситуация возникает, если дискретизация координат была проведена с шагом, сопоставимым либо превышающим дисперсию распределения по каждой координате.

Линеаризация системы уравнений для выбора оптимального начального приближения

Квадратичная форма $(\vec{x}_k - \vec{\mu})^T\, C\, (\vec{x}_k - \vec{\mu})$ представляется в виде:

$$\sum_{i=1}^{n} C_{ii}\left(x_k^{(i)} - \mu^{(i)}\right)^2 + \sum_{i=1}^{n}\sum_{j\neq i} C_{ij}\left(x_k^{(i)} - \mu^{(i)}\right)\left(x_k^{(j)} - \mu^{(j)}\right), \tag{2}$$

где C_{ii} имеет смысл масштаба дисперсии распределения по координате i (), а C_{ij} имеет смысл коэффициента взаимозависимости координат i и j. В реальных системах коэффициенты взаимозависимости обычно имеют значительно меньший вес, нежели дисперсия. Если C_{ij} можно пренебречь, то оценка C_{ii} получается из известных соотношений значений:

$$\frac{f_k}{A_0} = \exp\left(-\sum_{i=1}^{n} C_{ii}\left(x_k^{(i)} - \mu_0^{(i)}\right)^2\right), \tag{3}$$

куда поставляются полученные ранее начальные приближения по максимуму значения $\vec{\mu}_0$ и A_0.

Чтобы уточнить начальное приближение среднего $\vec{\mu}_0$, делается «нулевой шаг» – линеаризация системы уравнений $f_k = f(\vec{x}_k)$ и нахождение её решения каким-либо одношаговым методом, например, сингулярным разложением.

Для исключения квадратичного члена из системы уравнений поделим все уравнения на f_m:

$$\begin{aligned}\frac{f_k}{f_m} &= \frac{A\exp\left(-\sum_{i=1}^{n} C_{ii}\left(x_k^{(i)} - \mu^{(i)}\right)^2\right)}{A\exp\left(-\sum_{i=1}^{n} C_{ii}\left(x_m^{(i)} - \mu^{(i)}\right)^2\right)} \\ &= \exp\left(-\sum_{i=1}^{n} C_{ii}\left[\left(x_k^{(i)}\right)^2 - \left(x_m^{(i)}\right)^2 - 2\left(x_k^{(i)} - x_m^{(i)}\right)\mu^{(i)}\right]\right).\end{aligned} \tag{4}$$

Видно, что полученная система уравнений не зависит от высоты A, но в ней всё ещё присутствует масштаб дисперсии C_{ii}. Чтобы избавиться от него, введём нормированные координаты $u^{(i)} = x^{(i)}\sqrt{C_{ii}}$:

$$\frac{f_k}{f_m} = \exp\left(-\sum_{i=1}^{n}\left[\left(u_k^{(i)}\right)^2 - \left(u_m^{(i)}\right)^2 - 2\left(u_k^{(i)} - u_m^{(i)}\right)\eta^{(i)}\right]\right), \tag{5}$$

где $\eta^{(i)} = \mu^{(i)}\sqrt{C_{ii}}$.

Логарифмируем полученное выражение и получим систему $n-1$ уравнений, линейных относительно нормированного среднего $\vec{\eta}$:

$$\begin{aligned}&\ln f_k - \ln f_m = -\sum_{i=1}^{n}\left[\left(u_k^{(i)}\right)^2 - \left(u_m^{(i)}\right)^2 - 2\left(u_k^{(i)} - u_m^{(i)}\right)\eta^{(i)}\right]; \\ &2\sum_{i=1}^{n}\left(u_k^{(i)} - u_m^{(i)}\right)\eta^{(i)} = \ln f_k - \ln f_m + \sum_{i=1}^{n}\left[\left(u_k^{(i)}\right)^2 - \left(u_m^{(i)}\right)^2\right].\end{aligned} \tag{6}$$

Эта система имеет вид $U_{ik}\eta^{(i)} = v_k$, либо в матричном виде $U\vec{\eta} = \vec{v}$. Решение такой системы – оптимальное начальное приближение среднего $\vec{\eta}_0$, которое будет достаточно близко к истинному среднему $\vec{\eta}$, если полученное ранее приближение масштаба дисперсии достаточно близко к истинному C_{ii}. Стоит заметить, что такая линеаризация Гауссова купола возможна только при малости шума в значениях f_k, поскольку тогда его распределение приблизительно совпадает с распределением его логарифма.

Проверка метода

Вышеописанный метод был проверен на двумерном Гауссовом куполе:

$$f_k = f(x_k, y_k) = A\exp\left(-\frac{(x-\mu_x)^2 + \left(y-\mu_y\right)^2}{\sigma^2}\right), \tag{7}$$

значения которого были получены на квадратной сетке с шагом $d = l \cdot \sigma$, где $d \in [0{,}5; 1{,}5]$. Столь редко дискретизованные значения не позволяют точно оценить масштаб дисперсии σ, если среднее $\vec{\mu} = (\mu_x, \mu_y)$ не находится рядом с одной из точек сетки. Следовательно, нужно воспользоваться априорным знаниями о распределении: $\sigma_0 \approx d, C_{ii} \approx 1/d^2$.

Исходные данные для поиска генерировались со случайными значениями расположения среднего $\vec{\mu}$, высоты A и дисперсии σ. Выбрана решётка размером 7 × 7 узлов, в которой $\vec{\mu}$ находится в пределах одной ячейки рядом с центром. Таким образом, все точки находятся не далее 3σ от среднего, что соответствует 98,89% от всех значений распределения [1].

К полученным значениям функции был применён аддитивный белый гауссов шум с отношением сигнал-шум 30 дБ, симулирующий неточности измерений. Поскольку точки, в которых значения функции сопоставимы с амплитудой шума, вносят сильные искажения в исходные данные, перед применением алгоритма были удалены точки, значения которых оказались ниже 5% от максимального измеренного значения.

Оптимизация функционала квадратичного отклонения выполнялась по алгоритму Левенберга-Марквардта [2,3], реализованному в библиотеке Eigen [4], с целевой точностью 10^{-6} по смещению и градиенту. В качестве начального приближения для контрольного эксперимента было выбрано положение точки сетки с максимальным измеренным значением функции. В итоге было получено отклонение вычисленного положения среднего от истинного на ~0,58%·d в среднем или на ~2.14%·d в худшем случае (см. рисунок 1(а)). Для достижения минимума функционала потребовалось в среднем 5,6 шагов, в худшем случае 9 шагов.

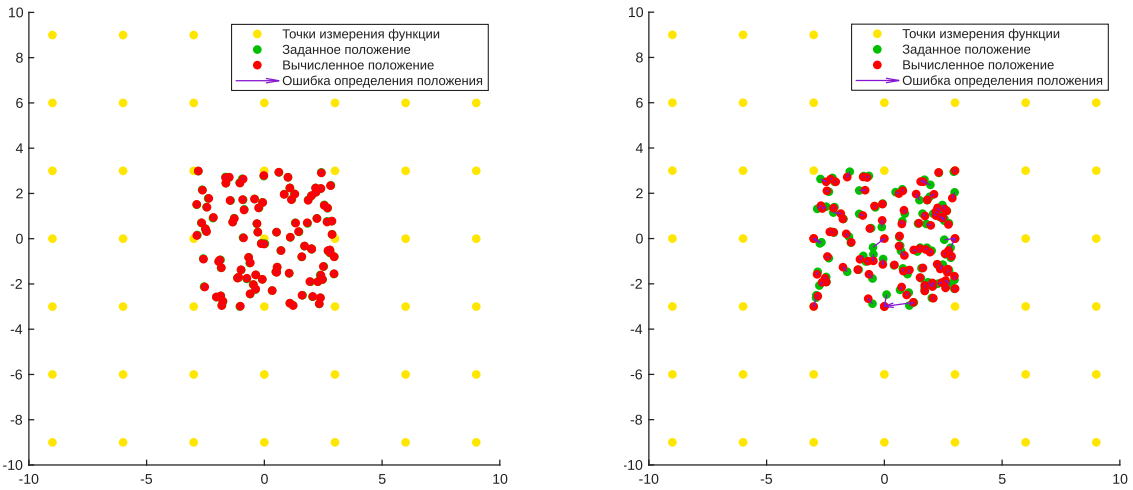


Рисунок 1 – Истинные (зелёный) и вычисленные (красный) средние значения распределений:
(а) после оптимизации функционала квадратичного отклонения;
(б) начальные приближения, полученные из системы уравнений.

Вычисленное оптимальное начальное приближение среднего $\vec{\mu}$ имеет отклонение от истинного на ~5,33%·d в среднем (см. рисунок 1(б)), что сопоставимо с худшим результатом оптимизации. В худшем случае отклонение составило ~47.98%·d. Для сравнения, при выборе начального приближения по максимальному значению функции среднее отклонение составило ~38.38%·d, а в худшем случае – ~70,54%. Последующая оптимизация сходится к тем же положениям среднего, но за меньшее количество шагов – 4,2 в среднем или 5 в худшем случае.

Выводы

Предложен метод вычисления начального приближения среднего положения Гауссова купола, основанный на линеаризации системы уравнений в точках с известными значениями функции. Для его применения необходимо сделать два допущения: малости шума и низкой взаимозависимости координат. Численный эксперимент показал эффективность предложенного метода на редко дискретизованной сетке при наличии небольшого шума в измеренных значениях. Вычисленное оптимальное начальное приближение среднего положения в среднем имеет в 7,2 раз лучшую точность, чем начальное приближение по максимальному значению. При применении нового начального приближения получено ускорение процедуры оптимизации на 25-45%. С учётом дополнительных вычислений перед выполнением процедуры оптимизации предложенный метод обеспечил средний прирост производительности до 33% на модельных данных.

Литература

1. Wang B., Shi W., Miao Z. Confidence Analysis of Standard Deviational Ellipse and Its Extension into Higher Dimensional Euclidean Space // PLOS ONE. Public Library of Science, 2015. Т.10, №3. С. 1-17.
2. Levenberg K. A Method for the Solution of Certain Non-Linear Problems in Least Squares // Quarterly of Applied Mathematics. 1944. Т.2, №2. С. 164-168.
3. Marquardt D. An Algorithm for Least Squares Estimation of Nonlinear Parameter // SIAM Journal on Applied Mathematics. 1963. Т.11, №2. С. 431-441.
4. Guennebaud G. Eigen: a c++ linear algebra library // Libraries for scientific computing. Ecole Polytechnique, 2013. [Электронный ресурс] URL : https://libeigen.gitlab.io/papers/eigen_aristotle_may_2013.pdf.

OPTIMAL INITIAL APPROXIMATION FOR NUMERICAL PARAMETER SEARCH OF A TWO-DIMENSIONAL GAUSSIAN BELL FUNCTION

Nikitin O.V.[1], Ovchinnikov P.E.[1]

[1] Lobachevsky State University of Nizhny Novgorod

The problem of numerical search for the mean of a two-dimensional Gaussian bell function from measured values at specific sampling grid points is considered. Its least-squares solution is found using iterative optimization. To increase computation speed, a new method for calculating the optimal initial approximation of the mean location, based on linearization of the equation system, has been developed. Utilizing this method on a wide-step sampling grid, a decrease of up to 33% in algorithm computation time was observed, when compared with the highest-value initial approximation.

ОЦЕНКА ПАРАМЕТРОВ ПОЛИГАРМОНИЧЕСКОГО СИГНАЛА НА ОСНОВЕ МИНИМИЗАЦИИ АТОМАРНОЙ НОРМЫ

студ. Лазько Е.В., доц. Маврычев Е.А.

Нижегородский государственный университет им. Н.И. Лобачевского
Нижегородский государственный технический университет им. Р.Е. Алексеева

В данной работе рассматривается задача спектрального анализа сигналов с повышенным разрешением. Спектральный анализ сигналов с использованием методов, основанных на восстановлении разреженных сигналов (sparse recovery) позволяет достичь существенного улучшения характеристик разрешения по сравнению с известными классическими алгоритмами. Одним из наиболее распространенных подходов разреженного восстановления является минимизация квадрата ошибки, регуляризированного нормой первого порядка вектора неизвестных параметров. Методы разреженного восстановления основываются на дискретном представлении неизвестного параметра, что позволяет сформировать конечное множество векторов, через линейную комбинацию которых представляется входной сигнал. Однако это условие противоречит практической постановке задачи спектрального анализа, в которой частоты гармоник принадлежат непрерывному множеству значений. В рассматриваемом подходе предполагается непрерывное множество параметров, а разреженность характеризуется атомарной нормой вектора сигнала. В работе представлен алгоритм, основанный на минимизации атомарной нормы, использующий выпуклое программирование с полуопределенными условиями. Приведены результаты математического моделирования, показывающие разрешающие способности и точности оценки частот спектральных составляющих по сравнению с хорошо известным алгоритмом MUSIC.

Введение

Методы спектрального анализа находят применения в различных областях обработки сигналов, включая радио- и гидролокацию, радиоастрономию, сейсмологию, биомедицину и т.д. Разрешающая способность классических алгоритмов спектрального анализа, таких как метод периодограмм или коррелограммный метод [1]-[2], ограничена пределом Релея, который определяется временем наблюдения. Однако, для многих задач спектрального анализа представляют интерес алгоритмы, позволяющие повысить разрешающую способность без увеличения времени наблюдения. Известно большое число методов спектрального анализа с повышенным разрешением. К ним относят метод минимальной дисперсии, метод MUSIC и другие алгоритмы, представленные в работах [1]-[3].

В обработке сигналов особое внимание получили методы разреженного восстановления сигналов (sparse signal processing) [4]-[6], в том числе и для задач спектрального анализа, так как эти методы позволяют значительно повысить разрешающую способность [7]-[9]. Применение методов разреженной обработки в спектральном анализе основывается на предположении, что спектр сигнала состоит из малого числа дискретных гармоник. Восстановление разреженных спектров основывается на поиске минимального числа гармонических составляющих и их комплексных амплитуд, при котором обеспечивается значение квадрата ошибки меньше заданной величины. Получаемая задача оптимизации является невыпуклой, не имеющей решения полиномиальной сложности [4]. В связи с этим было предложено большое количество методов решения задачи разреженной аппроксимации, которые позволяют найти какое-либо приближенное решение с приемлемой вычислительной сложностью. Наиболее эффективным является подход, основанный на регуляризации квадрата ошибки нормой первого порядка информационного вектора [9]. Такая задача в регрессионном анализе получала название LASSO [10], она является выпуклой и может быть решена методами выпуклого программирования [11].

Основным недостатком методов разреженной аппроксимации спектра является условие дискретности частот гармоник, которые должны соответствовать некоторому конечному множеству. При нарушении этого условия характеристики алгоритмов разреженного восстановления могут существенно снижаться. В данной работе представлен подход, основанный на использовании разреженности для непрерывного множества оцениваемых параметров. Такую разреженность характеризуют атомарной нормой, минимизация которой сводится к минимизации ранга Теплицевой неотрицательно определенной корреляционной матрицы полигармонического сигнала [12]-[14]. Далее целевая функция оптимизационной задачи аппроксимируется выпуклой функцией следа матрица и на основе ее решения оцениваются значения частот полигармонического сигнала. Результаты математического

моделирования демонстрируют существенно более высокую разрешающую способность предлагаемого алгоритма по сравнению с алгоритмом MUSIC.

Модель сигнала

Полагаем, что в входной сигнал $x(t) \in \mathbb{C}$ является комплексным полигармоническим, в котором присутствует конечное число гармонических составляющих, равное K. Рассмотрим задачу спектрального анализа, в которой по выборке дискретных комплексных отсчетов входного сигнала $x_1, x_1, \dots, x_L \in \mathbb{C}$, необходимо определить спектральные параметры сигнала: частоты гармоник и их мощности, где $x_l = x(\Delta t \cdot l)$, Δt - интервал дискретизации, $l = 1, \dots, L$ – дискретное время, L - длина выборки. Наблюдаемый входной сигнал запишем в виде

$$x_l = \sum_{k=1}^{K} a_k e^{j\omega_k l} + v_l, \;\; l = 1, \dots, L, \tag{1}$$

где a_k – комплексная амплитуда k-ой гармоники, ω_k – круговая частота k-ой гармоники, v_l – аддитивный белый шум с нулевым математически ожиданием и единичной дисперсией. В векторном виде модель входного сигнала (1) перепишем как

$$\mathbf{x} = \mathbf{z} + \mathbf{v} = \sum_{k=1}^{K} a_k \mathbf{s}(\omega_k) + \mathbf{v}, \tag{2}$$

где $\mathbf{x} = [x_1, x_2, \dots, x_L]^T \in \mathbb{C}^L$ – вектор входного сигнала, $\mathbf{z} = [z_1, z_2, \dots, z_L]^T \in \mathbb{C}^L$ – вектор полигармонического сигнала без шума, $\mathbf{v} = [v_1, v_2, \dots, v_L]^T \in \mathbb{C}^L$ – вектор собственного шума, $\mathbf{s}(\omega) = \left[1, e^{j\omega}, \dots, e^{j\omega(L-1)}\right]^T \in \mathbb{C}^L$ – вектор комплексного гармонического сигнала единичной амплитуды на частоте ω, $[.]^T$ - транспонирование вектора или матрицы.

Введем в рассмотрение корреляционную матрицу полигармонического сигнала $\mathbf{R}_z = \mathbb{E}\{\mathbf{z}\mathbf{z}^H\}$ и корреляционную матрицу входного сигнала $\mathbf{R}_x = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$, имеющие размерности $L \times L$, где $\mathbb{E}\{.\}$ - математическое ожидание, $[.]^H$ – эрмитово сопряжение. Будем полагать статистическую независимость комплексных амплитуд полигармонического сигнала. Тогда корреляционная матрица полигармонического сигнала представляется в виде

$$\mathbf{R}_z = \sum_{k=1}^{K} p_k \mathbf{s}(\omega_k)\mathbf{s}^H(\omega_k) = \mathbf{S}\mathbf{P}\mathbf{S}^H, \tag{3}$$

где $\mathbf{S} = [\mathbf{s}(\omega_1), \mathbf{s}(\omega_2), \dots, \mathbf{s}(\omega_K)]^T \in \mathbb{C}^{L \times K}$ - матрица комплексных гармоник, присутствующих во входном сигнале, $\mathbf{P} = \mathrm{diag}\{p_1, p_2, \dots, p_K\}$ - диагональная матрица, содержащая средние мощности гармоник, т.е. $p_k = \mathbb{E}\{|a_k|^2\}$.

Рассмотрим свойства матрицы $\mathbf{R}_z$. Ранг этой матрицы не превышает K, а если гармоники имеют отличные друг от друга частоты, то ранг равен K, т.е. $\mathrm{rank}\{\mathbf{R}_z\} = K$. Кроме того, все собственные числа матрицы $\mathbf{R}_z$ являются действительными и больше или равны нулю. Это означает что матрица $\mathbf{R}_z$ - положительно полуопределенная или неотрицательно определенная.

Матрица $\mathbf{s}(\omega)\mathbf{s}^H(\omega)$ для $\forall\omega$ является Теплицевой матрицей, так как (n, m)-ый элемент равен $e^{j\omega(n-m)}$. Тогда матрица $\mathbf{R}_z$, являющаяся суммой Теплицевых матриц, также является Теплицевой матрицей и ее можно представить в виде

$$\mathbf{R}_z = \mathbf{T}(\mathbf{u}) = \begin{bmatrix} u_1 & u_2 & \dots & u_L \\ u_2^* & u_1 & \dots & u_{L-1} \\ & & \ddots & \\ u_L^* & \dots & u_2^* & u_1 \end{bmatrix}, \tag{4}$$

где $\mathbf{u} = [u_1, u_2, \dots, u_L]^T \in \mathbb{C}^L$ - вектор автокорреляции полигармонического сигнала.

Таким образом, матрица $\mathbf{R}_z$ является Теплицевой неотрицательно определенной матрицей с рангом равным K.

Учитывая, что собственный шум является белым и имеет нулевое математическое ожидание и дисперсию равную единице, корреляционная матрица входного сигнала будет иметь вид

$$\mathbf{R}_x = \mathbf{S}\mathbf{P}\mathbf{S}^H + \mathbf{I}_L, \tag{5}$$

где $\mathbf{I}_L$ - единичная матрица размерности $L \times L$.

Атомарная норма

Если сигнал $x(t)$ представляет собой суперпозицию K гармоник, то его спектр при отсутствии шума будет являться разреженным, т.е. содержать только K ненулевых значений на частотах $\omega_1, \omega_2, \ldots, \omega_K$. В работах [7], [9] предполагается модель сигнала, в котором частоты гармоник принадлежат дискретному множеству значений, что является непрактичным ограничением. Рассмотрим сигналы, у которых значения частот принадлежат непрерывному множеству.

Пусть дискретный полигармонический сигнал представляется в виде

$$\mathbf{z} = \sum_{k=1}^{K} a_k \mathbf{s}_k. \tag{6}$$

Этот сигнал состоит из весовой суперпозиции векторов $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K \in S$, в которой весами являются комплексные амплитуды $a_1, a_2, \ldots, a_K$. Другими словами, вектор $\mathbf{z}$ является линейной комбинацией векторов принадлежащих бесконечному множеству $S := \{\mathbf{s}(\omega)\}_{\omega \in [0, 2\pi]}$. Любой вектор из множества S является атомом, а любая линейная комбинация атомов из S образует полигармонический сигнал.

По аналогии с векторной нормой нулевого порядка, которая равна числу ненулевых компонент вектора, и характеризующая разреженность вектора, будем рассматривать атомарную норму $\|\mathbf{z}\|_{\mathcal{A}}$, которая равна числу гармоник, присутсвующих в сигнале. Математическое определение атомрной нормы представляется как нижняя граница множества всевозмжного числа гармонических состаляющих, которыми можно аппроксимировать вектор полигармонического сигнала $\mathbf{z}$, т.е. атомарную норму определим в виде

$$\|\mathbf{z}\|_{\mathcal{A}} := \inf\left\{ \mathcal{K}: \mathbf{z} = \sum_{k=1}^{\mathcal{K}} a_k \mathbf{s}_k, \qquad \mathbf{s}_k \in S, a_k \in \mathbb{C} \right\}, \tag{7}$$

где $\inf\{.\}$ – нижняя граница множества, $\mathcal{K}$ - множество, включающее всевозможное число гармоник, аппроксимирующих вектор $\mathbf{z}$.

Минимизация атомарной нормы

Рассмотрим постановку задачи спектрального разложения на основе минимизации разреженности спектра или минимизации атомарной нормы вектора полигармонического сигнала $\mathbf{z}$. Из определения атомарной нормы следует, что ее значение равно рангу матрицы $\mathbf{R}_z$, так как он равен числу линейно независимых гармоник, присутствующих в сигнале, т.е.

$$\|\mathbf{z}\|_{\mathcal{A}} = \operatorname{rank}\{\mathbf{R}_z\}. \tag{8}$$

Тогда задача минимизации атомарной нормы эквивалента поиску такого вектора $\mathbf{u}$, при котором достигается минимум ранга Теплицевой матрицы $\mathbf{T}(\mathbf{u})$ при условии, что эта матрица является неотрицательно определенной

$$\min_{\mathbf{u}} \operatorname{rank}\{\mathbf{T}(\mathbf{u})\}, \qquad \boldsymbol{s.t.} \qquad \mathbf{T}(\mathbf{u}) \succ \mathbf{0}. \tag{9}$$

Однако, в выражении (9) не хватает условий, которые бы связывали Теплицеву матрицу $\mathbf{T}(\mathbf{u})$ с вектором сигнала $\mathbf{z}$. Рассмотрим аналитическое выражение для компонент Теплицевой корреляционной матрицы $\mathbf{R}_z = \mathbf{T}(\mathbf{u})$. Элемент u_l записывается как

$$u_l = \sum_{k=1}^{K} p_k e^{j\omega_k(l-1)}, \qquad l = 1, \ldots, L. \tag{10}$$

Из (10) следует, что вектор $\mathbf{u}$ является линейной комбинацией векторов $\mathbf{s}(\omega_1), \mathbf{s}(\omega_2), \ldots, \mathbf{s}(\omega_K)$, т.е. тех же самых векторов, что и вектор $\mathbf{z}$, исходя из (1), (2). Тогда если корреляционную матрицу дополнить вектором $\mathbf{z}$ в виде

$$\widetilde{\mathbf{R}}_z = \begin{bmatrix} x & \mathbf{z}^H \\ \mathbf{z} & \mathbf{T}(\mathbf{u}) \end{bmatrix}, \quad (11)$$

то дополненная матрица (11) имеет тот же базис собственных векторов с увеличенной размерностью, что и матрица $\mathbf{R}_z$, а также имеет ранг равный числу линейно независимых гармонических составляющих и является неотрицательно определенной. В матрице (11) $x \in \mathbb{R}$, $x>0$ – свободная переменная.

Таким образом, если вместо $\mathbf{T}(\mathbf{u})$ использовать дополненную матрицу (11), то получим оптимизационную задачу относительно вектора $\mathbf{u}$ при известном полигармоническом сигнале $\mathbf{z}$. Эту задачу оптимизации запишем в виде

$$\min_{x,\mathbf{u}} \operatorname{rank}\left\{\begin{bmatrix} x & \mathbf{z}^H \\ \mathbf{z} & \mathbf{T}(\mathbf{u}) \end{bmatrix}\right\}, \ \boldsymbol{s.t.} \quad \begin{bmatrix} x & \mathbf{z}^H \\ \mathbf{z} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \succ \mathbf{0}. \quad (12)$$

Ранг матрицы является невыпуклой функцией и решение подобной оптимизационной задачи представляется крайне сложным с вычислительной точки зрения. Метод решения задачи (12) будет обладать экспоненциальной вычислительной сложностью. В [14] предлагается аппроксимация целевой функции (12) с помощью следа матрицы $Tr\{\widetilde{\mathbf{R}}_z\}$, что приводит к задачи выпуклого программирования

$$\min_{x,\mathbf{u}}(x + Nu_1), \ \boldsymbol{s.t.} \quad \begin{bmatrix} x & \mathbf{z}^H \\ \mathbf{z} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \succ \mathbf{0}. \quad (13)$$

В оптимизационной задаче (13) используется «чистый» полигармонический сигнал $\mathbf{z}$ без шума, в то время как на практике имеется сигнал $\mathbf{x}$. В этом случае к целевой функции (13) добавляется квадрат ошибки $\|\mathbf{z}-\mathbf{x}\|_2^2$, где $\mathbf{z}$ рассматривается как неизвестный вектор. Тогда $(x + Nu_1)$ играет роль регуляризирующей функции. Окончательно оптимизационная задача записывается в виде

$$\min_{x,\mathbf{u},\mathbf{z}}\{\lambda(x + Nu_1) + \|\mathbf{z}-\mathbf{x}\|_2^2\}, \ \boldsymbol{s.t.} \quad \begin{bmatrix} x & \mathbf{z}^H \\ \mathbf{z} & \mathbf{T}(\mathbf{u}) \end{bmatrix} \succ \mathbf{0}. \quad (14)$$

Оптимизационная задача (14) является выпуклой и относится к классу задач с полуопределенными условиями, которая эффективно решается методом внутренней точки [11].

После того как получено решение задачи (14), оценка частот может быть реализована на основе подхода подобному алгоритму ESPRIT [15]. Для полученной Теплицевой матрицы используем разложение Холевского

$$\mathbf{T}(\mathbf{u}) = \mathbf{V}\mathbf{V}^H. \quad (15)$$

Рассмотрим матрицы $\mathbf{V}_{-1}$ и $\mathbf{V}_{-L}$ размерности $(L-\mathbf{1}) \times L$, полученные из матрицы $\mathbf{V}$ путем удаления первой строки и последней строки, соответственно. Эти матрицы связаны между собой через преобразование унитарной матрицей $\mathbf{Q}$, которое записывается как $\mathbf{V}_{-1} = \mathbf{Q}\mathbf{V}_{-L}$. Тогда получим

$$\mathbf{Q} = \left(\mathbf{V}_{-L}\mathbf{V}_{-L}^H\right)^{-1}\mathbf{V}_{-L}^H\mathbf{V}_{-1}. \quad (16)$$

Матрица $\mathbf{Q}$ является матрицей сдвига, ее собственные числа имеют единичную амплитуду, а фазы определяются частотами гармоник. Матрицу сдвига можно разложить в базис собственных векторов

$$\mathbf{Q} = \mathbf{U}\widetilde{\mathbf{Q}}\mathbf{U}^H, \quad (17)$$

где $\widetilde{\mathbf{Q}} = \operatorname{diag}\{q_1, q_2, \ldots, q_K\}$ - диагональная матрица собственных чисел, которые равны

$$q_k = e^{j\omega_k}, \ \ k = 1, \ldots, K. \quad (18)$$

Исходя из (18) определяются оценки частот гармоник.

Вычислительная сложность алгоритма минимизации атомарной нормы определяется главным образом сложностью решения задачи (14). Асимптотическая сложность решения оптимизационной задачи полуопределенного программирования оценивается как $O(L+1)^{3,5}$.

Результаты моделирования

Для оценки эффективности предлагаемого алгоритма проводится статистическое моделирование с усреднением по 200 реализациям, в котором получены среднеквадратические ошибки оценок частоты. Метод атомарной нормы (ANM) сравнивается с алгоритмом MUSIC. Рассматривается полигармонический сигнал, состоящий из двух гармоник равными амплитудами и случайными начальными фазами. Частоты гармоник равны $\omega_1 = 2\pi \cdot 0.1,\ \omega_2 = 2\pi \cdot (0.1 + \Delta f)$. Число выборок входного процесса $L = 50$.

На рис. 1 построена зависимость среднеквадратической ошибки (RMSE) от отношения сигнал/шум (SNR), при этом $\Delta f = 0.01$. На рис. 2 представлена среднеквадратическая ошибка в зависимости от разноса частот Δf при постоянном отношении сигнал/шум, равным 10 дБ. Представленные рисунки показывают, что алгоритм минимизации атомарной нормы дает существенно лучшие пороговые характеристики. Алгоритм MUSIC достигает ошибки сравнимой с алгоритмом атомарной нормы при ОСШ на 12 дБ больше. Однако, предлагаемый алгоритм существенно уступает в точности оценивания при больших значениях ОСШ.

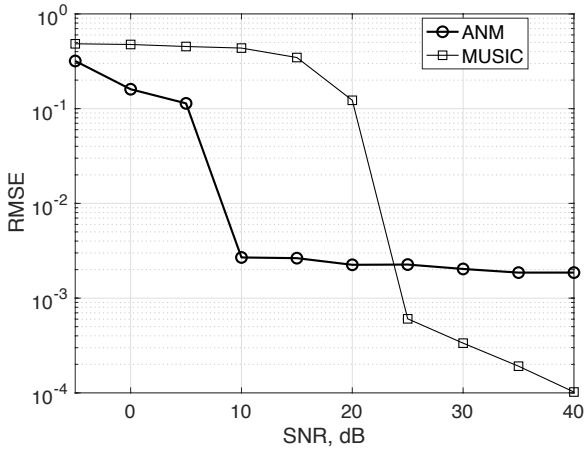


Рисунок 1 – Зависимость среднеквадратической ошибки оценки частоты от ОСШ

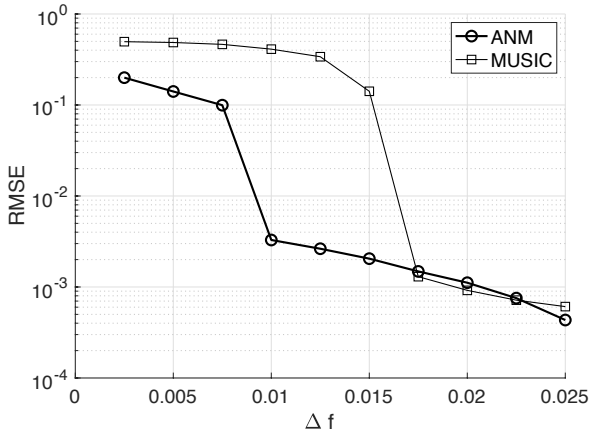


Рисунок 2 – Зависимость среднеквадратической ошибки оценки частоты от разноса частот

Заключение

Таким образом, в данной работе предложен алгоритм оценки частот полигармонического сигнала на основе минимизации атомарной нормы вектора сигнала. Данный алгоритм выполняет оценку корреляционной функции полигармонического сигнала на основе решения выпуклой оптимизационной задачи с полуопределенными условиями. Далее, аналогично алгоритму ESPRIT, вычисляется унитарная матрица сдвига, собственные числа которой непосредственно связаны с частотами гармоник. Результаты математического моделирования показывают, что метод минимизации атомарной нормы имеет существенно лучшую ошибку по сравнению с MUSIC при малом ОСШ (рис. 1) и при малом разносе частот (рис. 2), где метод MUSIC еще не разрешает две гармоники. Т.е. метод атомарной нормы имеет существенно лучшее пороговое разрешение.

Литература

1. С.Л. Марпл. Цифровой спектральный анализ и его приложения // М.: Мир, 1990. 584 с.
2. P. Stoica, R. Moses. Spectral analysis of signals // Prentice Hall, NJ, 2005. 427 p.
3. Б.И. Шахтарин, В.А. Ковригин. Методы спектрального оценивания случайных процессов // М: Горячая линия – Телеком, 2023. 256 с.
4. S. Foucart, H. Rauhut. A mathematical introduction to compressive sensing // Birkhäuser, 2013. – 625 p.
5. D. Donoho. Compressed sensing // IEEE Transactions on Information Theory, 2006, Vol. 52, No. 4. - Pp. 1289-1306.
6. D.L. Donoho, X. Huo. Uncertainty principles and ideal atomic decomposition // IEEE Transactions on Information Theory, 2001, Vol. 47, No. 7. - Pp. 2845-2862.
7. I.F. Gorodnitsky, B.D. Rao. Sparse signal reconstruction from limited data using FOCUSS: a reweighted minimum norm algorithm // IEEE Trans. on Signal Processing, 1997, Vol. 45, No. 3. - Pp. 600–616.
8. M. Duarte, R. Baraniuk. Spectral compressive sensing // Applied and Computational Harmonic Analysis, 2013, Vol. 35, No 1. – Pp. 111-129.
9. А.П. Пехтерев, Е.А. Маврычев. Параллельная реализация итерационного алгоритма спектрального анализа на основе регуляризации нормой первого порядка // Радиолокация, навигация, связь: сборник трудов XXVIII Международной научно-технической, посвященной памяти конференции Б.Я. Осипова, г. Воронеж, 27-29 сентября 2022, Т. 1, С. 213-224.
10. R. Tibshirani. Regression shrinkage and selection via the Lasso // Journal of the Royal Statistical Society Series B, 1996, Vol. 58, No. 1. - Pp. 267–288.
11. S. Boyd, L. Vandenberghe. Convex optimization. Cambridge University Press, 2004. – 716 p.
12. B. N. Bhaskar, G. Tang, and B. Recht. Atomic norm denoising with applications to line spectral estimation // IEEE Transactions on Signal Processing, vol. 61, no. 23, pp. 5987–5999, 2013.
13. Z. Yang and L. Xie. On gridless sparse methods for line spectral estimation from complete and incomplete data // IEEE Trans. on Signal Processing, 2015, Vol. 63, No. 12. - Pp. 3139–3153.
14. Z. Yang, J. Li, P. Stoica, L. Xie. Sparse Methods for direction-of-arrival estimation // Academic Press Library in Signal Processing, Vol. 7, Pp. 509-581.
15. R. Roy, T. Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques // IEEE Trans. on Signal Processing, 1989, Vol. 37, No. 7. - Pp. 984–995.

PARAMETERS ESTIMATION OF POLYHARMONIC SIGNAL BASED ON ATOMIC NORM MINIMIZATION

Lazko E.V.[1], Mavrychev E.A.[2]

[1] Nizhny Novgorod State University n.a. N.I. Lobachevsky
[2] Nizhny Novgorod State Technical University n.a. R.E. Alekseev

This paper considers the problem of spectral analysis of signals with increased resolution. Spectral analysis of signals using methods based on sparse signal recovery allows for a significant improvement in resolution characteristics compared to well-known classical algorithms. One of the most common approaches to sparse recovery is minimizing the squared error regularized by the first-order norm of the unknown parameter vector. Sparse recovery methods are based on a discrete representation of the unknown parameter, which allows for the formation of a finite set of vectors through a linear combination of which the input signal is represented. However, this condition contradicts the practical formulation of the spectral analysis problem, in which the harmonic frequencies belong to a continuous set of values. The approach under consideration assumes a continuous set of parameters, and sparsity is characterized by the atomic norm of the signal vector. This paper presents an algorithm based on minimizing the atomic norm using convex programming with semidefinite conditions. The results of mathematical modeling are presented, demonstrating the resolution and accuracy of estimating the frequencies of spectral components compared to the well-known MUSIC algorithm.

АНАЛИЗ РАДИОЛОКАЦИОННЫХ ИЗОБРАЖЕНИЙ НАДВОДНЫХ ЦЕЛЕЙ НА ФОНЕ МОРСКОЙ ПОВЕРХНОСТИ НА ОСНОВЕ ИХ ПОЛИГОНАЛЬНЫХ И МНОГОТОЧЕЧНЫХ МОДЕЛЕЙ

студ. Зельницкий Н.А., проф. Сучков В.Б., ст. преп. Перов А.Ю.

МГТУ им. Н.Э. Баумана

Рассматривается задача формирования и анализа радиолокационных изображений надводных целей при наблюдении их на фоне морской поверхности. Для задач распознавания необходимы большие репрезентативные выборки изображений, однако получение высокодетализированных портретов по полигональным моделям требует значительных вычислительных затрат. Предлагается подход, сочетающий полигональное описание геометрии цели и ее многоточечное представление рассеивающих центров, с последующим синтезом изображений методом обратной проекции. Морская поверхность представляет собой подстилающую поверхность, влияющую на контраст и структуру портрета; исследуется, как изменение разрешающей способности по дальности и азимуту отражается на контрасте конструктивных элементов и устойчивости признаков на фоне морской поверхности. В среде MATLAB реализован комплекс имитационного моделирования, включающий обработку отражённых сигналов и построение портретов для набора ракурсов. Показано, что многоточечная модель обеспечивает сопоставимую информативность с полигональной при существенном ускорении массового формирования данных и позволяет обосновывать требования к разрешению при обнаружении и классификации надводных объектов.

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации по государственному заданию № FSFN-2025-0014.

Принцип формирования высококонтрастных радиолокационных изображений

Для формирования высококонтрастного радиолокационного изображения (радиопортрета) эсминца Zumwalt на фоне морской поверхности с волнением 2–3 балла используется модель отражённого сигнала, рассчитанная на основе полигонального представления модели. Поверхность корабля задаётся набором плоских треугольников и острых кромок, а характеристики рассеяния определяются с применением физической теории дифракции и геометрической оптики [1].

Точность синтеза радиопортрета напрямую зависит от степени детализации сетки: чем больше количество полигонов модели, тем корректнее описывается ее геометрия, однако тем выше вычислительные затраты. Для корабля класса Zumwalt расчёт отражённых сигналов по полной полигональной модели (рисунок 1) при множестве ракурсов оказывается слишком ресурсоёмким, особенно при учёте морского фона.

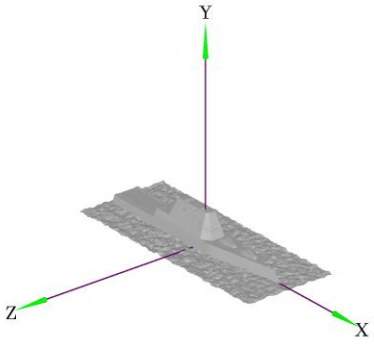


Рисунок 1 – Полигональная модель корабля Zumvalt на фоне морской поверхности

Поэтому полигональная модель используется как базовая для построения многоточечной модели — набора доминирующих рассеивающих центров, сохраняющих ключевые особенности рассеяния корабля, но позволяющих существенно ускорить массовое формирование радиопортретов. Моделирование выполняется для РЛС X-диапазона - 10 ГГц; характерный шаг дискретизации в проекционном объёме выбирается порядка 1 м, что обеспечивает компромисс между детализацией радиопортрета и скоростью вычислений.

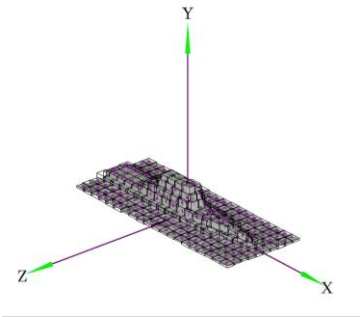


Рисунок 2 – Многоточечная модель корабля Zumvalt на фоне морской поверхности

Радиолокационный портрет, формируемый радиолокатором с инверсным синтезированием апертуры, можно рассматривать как проекцию дальностного и поперечного профилей цели на плоскость изображения. Эта плоскость задаётся опорной точкой (как правило, ею выбирают центр масс объекта) и направляющим вектором, в данном исследовании плоскость проекции была выбрана плоскость параллельная плоскости ZX. Отражённый сигнал накапливается при зондировании ЛЧМ-сигналом в пределах конечной полосы частот $\Delta f = f_{макс} - f_{мин}$, а также в заданном диапазоне азимутальных ракурсов $\alpha = \alpha_{мин}, \ldots, \alpha_{ц}, \ldots, \alpha_{макс}$.

Накопление откликов от точечного отражателя выполняется в пространственно-частотной области, задаваемой волновым числом $k = 2\pi/\lambda$ и углом визирования α. Поскольку и полоса частот Δf, и сектор синтезируемой апертуры $\Delta\alpha = \alpha_{макс} - \alpha_{мин}$ являются конечными, отображение накопленных данных на плоскость изображения приводит к формированию неравномерной сетки в пространстве пространственных частот (k_x, k_z). При достаточно малых значениях Δf и $\Delta\alpha$ эта сетка становится близкой к линейной и почти равномерной, что упрощает дальнейшее построение радиолокационного изображения.

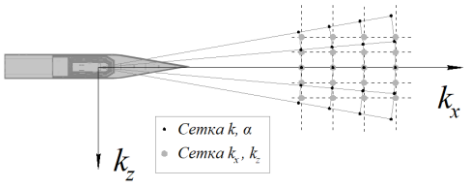


Рисунок 3 – Преобразование пространственной сетки (k, α) в пространственную сетку (k_x, k_z)

В силу допущения линейности пространственно-частотной области (k, α) для малых значений полосы зондирующего сигнала и угловых интервалов СА (синтезирования апертуры), представляется возможным применение процедуры обратного преобразования Фурье для отображения отражённого сигнала из пространственно-частотной области (k, α) в пространственную сетку

Корректный выбор шага пространственной сетки определяется требуемыми разрешениями по двум осям изображения — Δx, Δz. Эти значения следует задавать не крупнее характерного размера проекционного параллелепипеда многоточечной модели, чтобы радиолокационный портрет сохранял необходимую детализацию объекта локации.

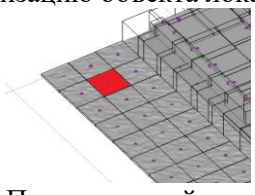


Рисунок 4 – Проекционный параллелепипед

При синтезе высококонтрастного радиолокационного портрета для выбранного ракурса входной сигнал РЛС рассчитывался на основе эффективных площадей рассеяния, полученных на этапе построения многоточечной модели. Комплексный коэффициент отражения каждого локального отражателя многоточечной модели определяется через параметры его рассеяния и

используется при формировании суммарного отклика, который затем преобразуется в радиолокационное изображение.

Комплексный коэффициент отражения локального отражателя многоточечной модели определяется из соотношения (1):

$$\dot{\rho}_k = \sqrt{\frac{\sigma_k}{4\pi}} exp\left(-j2\frac{2\pi}{\lambda}R_k\right), \quad (1)$$

где $\dot{\rho}_k$ – комплексный коэффициент отражения, σ_k – ЭПР k-й точки, которая вычисляется за счет интерполяционных соотношений для соседних значений диаграммы ЭПР, R_k – дальность от k-й точки модели до антенны РЛС.

Для полигональной модели цели выполняется численное моделирование полей рассеяния, на основе полученных результатов формируется матрица данных $\boldsymbol{M}$, представляющая собой совокупность измерений, каждый её элемент характеризуется комплексной амплитудой, и фазовым множителем, зависящим от расположения и параметров радара [2].

$$M_{f_i,\alpha_j} = \sum_{m=1}^{K} \dot{E}_m \exp\left(-j2\left(k_{ij}p_m\right)\right) \quad (2)$$

где: M_{f_i,α_j} - элемент матрицы данных с индексами i, j, $\dot{E}_m$ - поле рассеяния m - го отражателя, определяемые по выражениям из [2] полученные при использовании методов физической оптики и методов физической теории дифракции, $\boldsymbol{k}_{ij} = k_i\{cos(\alpha_j), sin(\alpha_j)\}$ - волновой вектор для комбинации f_i, α_j, $k_i = \frac{2\pi f_i}{c}$, $\boldsymbol{p}_m = \{x_m, z_m\}$ - направляющий вектор элементарного отражателя в системе координат цели. $K = K_\text{Т} + K_K + K_O$ - количество элементарных отражателей (треугольных элементов, острых кромок и переотражений соответственно) [1].

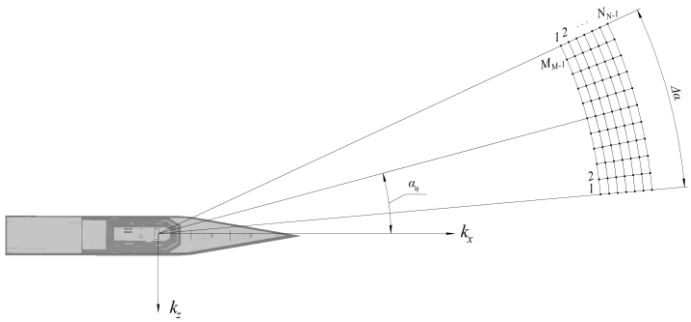


Рисунок 5 – Пространственная сетка частот и углов

Затем подготавливаются массивы частот и углов для расчетов (рисунок 5), которые задаются следующим образом:

$$f = \left[\left(f_0 - \frac{Ndf}{2}\right)\left(f_0 - \left(\frac{N}{2} - 1\right)df\right) \dots (f_0) \dots \left(f_0 + \frac{N}{2}df\right)\right]_{1\times N}; \quad (3)$$

$$\alpha = \left[\left(\alpha_ц - \frac{Md\alpha}{2}\right)\left(\alpha_ц - \left(\frac{M}{2} - 1\right)d\alpha\right) \dots (\alpha_ц) \dots \left(\alpha_ц + \frac{M}{2}d\alpha\right)\right]_{1\times M}; \quad (4)$$

где: M, N – количество шагов дискретизации пространственной сетки радиолокационного портрета.

Затем к полученной матрице применяется оконная функция:

$$M^{окон.}_{f_i,\alpha_j} = M_{f_i,\alpha_j} w_\alpha(j) w_f(i), \quad (5)$$

где: $w_\alpha(j)$ - оконная функция по углу, $w_f(i)$ - оконная функция по частоте.

При данном исследовании использовалось окно Кайзера с параметром $\beta = 12$.

Затем область построения изображения разбивается на точки (пиксели) согласно следующим выражениям:

$$x_u = -\frac{X}{2} + \frac{X}{M}u, \; u = 0 \dots M \quad (6)$$

$$z_u = -\frac{Z}{2} + \frac{Z}{N}v, \; v = 0 \dots N, \quad (7)$$

где: x_u, z_u - координаты точки на изображении в метрах, X, Z - размер изображения, M, N - число дискретов по осям x, z.

Тогда амплитуда поля в точке с координатами x_u, z_u определяется по соотношению (8).

$$ISAR(x_u,\ z_v) = \sum_{j=1}^{N_\alpha} \sum_{i=1}^{N_f} M_{f_i,\alpha_j}^{окон.}\, exp\left(j2\left(\boldsymbol{k}_{ij}\boldsymbol{r}_{uv}\right)\right), \quad (8)$$

где: N_α, N_f - число разбиений по углу места и частоте, $\boldsymbol{r}_{uv} = \{x_u, z_v\}$ - направляющий вектор на точку изображения с координатами x_u, z_u.

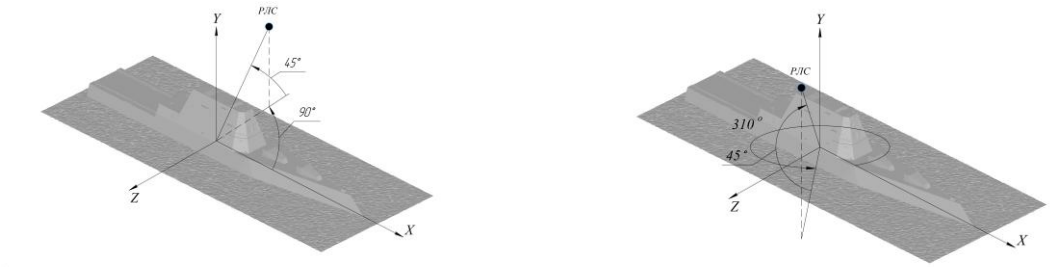
В случае построения радиолокационного изображения, $\dot{\rho}_m$ эквивалентно $\dot{E}_m$ для полигональной модели цели, тогда выражение (2) принимает следующий вид:

$$M_{f_i,\alpha_j} = \sum_{m=1}^{K} \dot{\rho}_m \exp\left(-j2\left(k_{ij}p_m\right)\right) \quad (9)$$

где: $\dot{\rho}_m$ - комплексный коэффициент отражения m - й блестящей точки многоточечной модели, $\boldsymbol{p}_m = \{x_m, z_m\}$ - направляющий вектор m - й блестящей точки.

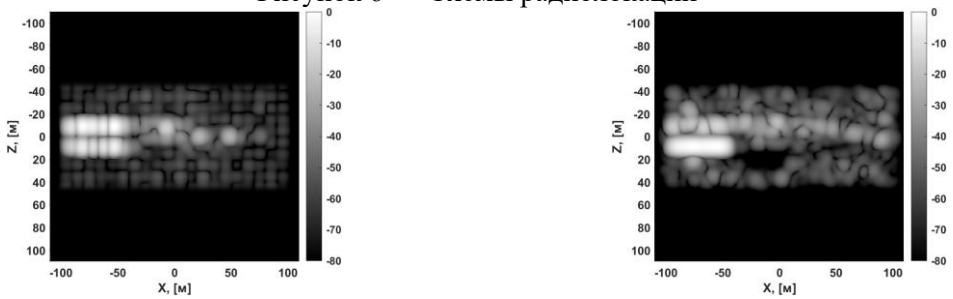
Таким образом, выражение (4) для многоточечной модели можно рассматривать как частный случай общего соотношения, применяемого при численном моделировании полей рассеяния полигональных объектов, где каждая элементарная грань или кромка представляется эквивалентным отражателем с комплексным коэффициентом $\dot{E}_m$ [3, 4].

Сравнение радиопортретов надводной цели при различных разрешениях



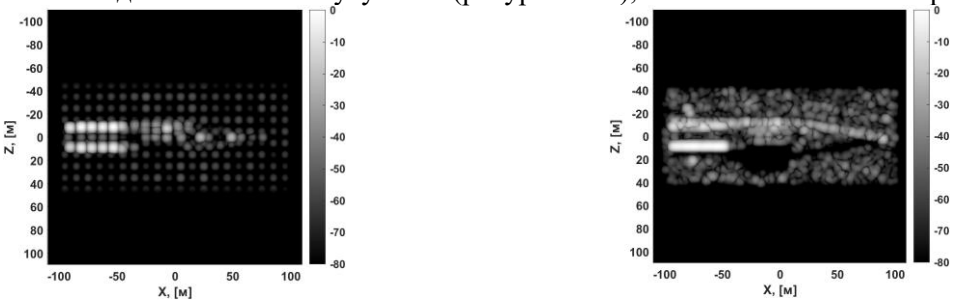
а) Ракурс 90 45; б) Ракурс 310 45

Рисунок 6 — Схемы радиолокации



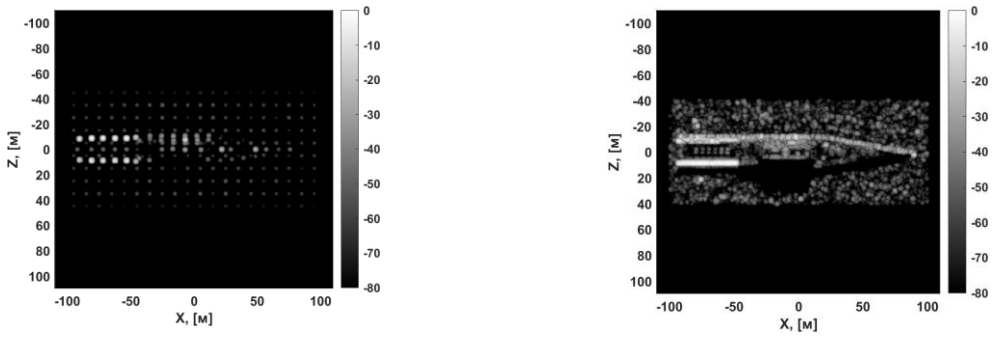
а) Многоточечная; б) полигональная

Рисунок 7 — Радиолокационное изображение, полученное для корабля Zumvalt при разрешении по дальности и азимуту в 3 м (ракурс 90 45), на согласованной поляризации



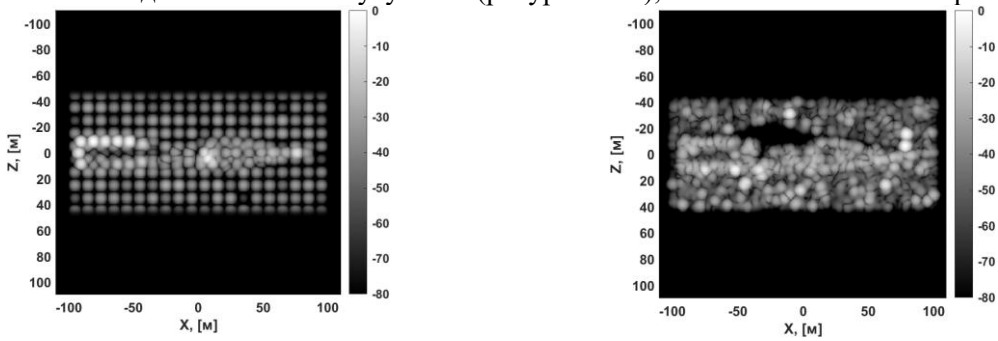
а) Многоточечная; б) полигональная

Рисунок 8 — Радиолокационное изображение, полученное для корабля Zumvalt при разрешении по дальности и азимуту в 1,5 м (ракурс 90 45), на согласованной поляризации



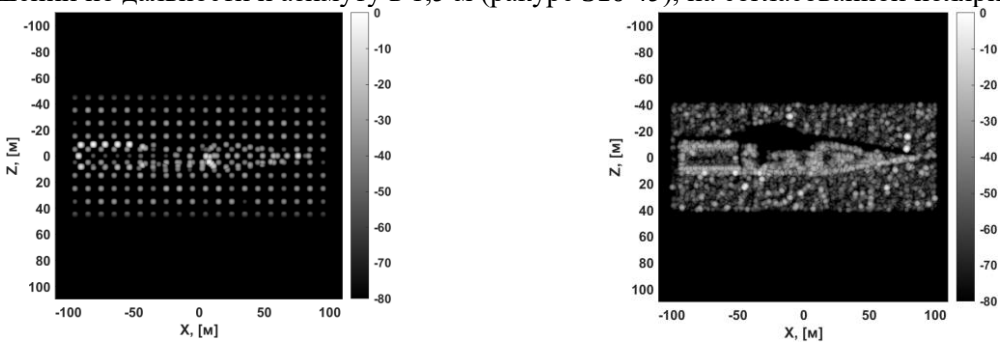
а) Многоточечная; б) полигональная

Рисунок 9 — Радиолокационное изображение, полученное для корабля Zumvalt при разрешении по дальности и азимуту в 1 м (ракурс 90 45), на согласованной поляризации



а) Многоточечная; б) полигональная

Рисунок 10 — Радиолокационное изображение, полученное для корабля Zumvalt при разрешении по дальности и азимуту в 1,5 м (ракурс 310 45), на согласованной поляризации



а) Многоточечная; б) полигональная

Рисунок 11 — Радиолокационное изображение, полученное для корабля Zumvalt при разрешении по дальности и азимуту в 1 м (ракурс 90 45), на согласованной поляризации

По представленным радиолокационным портретам (рис. 7–11) видно, что при уменьшении шага дискретизации и достижении пространственного разрешения по дальности и азимуту порядка 1–1,5 м обеспечивается требуемая детализация: стабилизируется пространственная структура портрета, а также сохраняются форма и взаимное расположение доминирующих яркостных максимумов, соответствующих ключевым конструктивным элементам цели. При более грубом разрешении (~3 м) наблюдается частичная потеря мелкомасштабных деталей и снижение контрастности отдельных элементов, однако общая конфигурация основных областей рассеяния остаётся воспроизводимой. Таким образом, достижение разрешения не хуже характерного размера проекционного объёма модели приводит к сопоставимой информативности изображений, сформированных по многоточечной и полигональной моделям, что подтверждает корректность выбора требуемого разрешения для задач обнаружения и классификации.

Заключение

В работе подтверждена эффективность предложенного алгоритма формирования радиолокационных радиопортретов надводной цели на фоне морской поверхности: синтезированные изображения характеризуются устойчивой пространственной структурой и воспроизводимостью доминирующих областей рассеяния при вариации ракурса наблюдения. Сопоставление результатов, полученных по полигональной и многоточечной моделям эсминца Zumwalt, показало, что при выборе пространственного разрешения изображения (по дальности и азимуту) не хуже характерного размера проекционного параллелепипеда многоточечной модели достигается высокая степень совпадения радиопортретов по форме и расположению основных яркостных максимумов. Это позволяет обоснованно применять многоточечную модель в качестве вычислительно более экономичного представления цели для формирования больших репрезентативных наборов радиопортретов, предназначенных для задач обнаружения и классификации надводных объектов.

Литература

1. Сучков, В. Б. Объектно-ориентированный метод определения комплексных коэффициентов отражения элементов полигональной модели объекта локации / В. Б. Сучков // Системы и средства связи, телевидения и радиовещания. – 2013. – № 1-2. – С. 159-165. – EDN RDVJQD.

2. Сучков, В. Б. Метод определения входных сигналов бортовых систем ближней радиолокации от объектов сложной формы на основе использования их полигональных и многоточечных моделей / В. Б. Сучков // Спецтехника и связь. – 2013. – № 3. – С. 25-31. – EDN QCLZFZ.

3. Сучков В. Б. Методика создания многоточечной модели аэродинамической цели для определения входных сигналов бортовых радиолокационных датчиков //Электромагнитные волны и электронные системы. – 2013. – Т. 18. – №. 6. – С. 45-53.

4. Borzov A. B. et al. Mathematical modeling and simulation of the input signals of short-range radar systems //Journal of Communications Technology and Electronics. – 2014. – V. 59. – P. 1356-1368.

ANALYSIS OF RADAR IMAGES OF SURFACE TARGETS AGAINST THE SEA-SURFACE BACKGROUND BASED ON THEIR POLYGONAL AND MULTIPOINT MODELS

Suchkov V.B.[1], Zelnitskiy N.A.[1], Perov A.U.

[1] Bauman Moscow State Technical University

The paper addresses the problem of generating and analyzing radar signatures of surface targets observed against the sea surface background. Large representative image datasets are required for recognition tasks; however, producing highly detailed signatures from polygonal target models entails significant computational cost. An approach is proposed that combines a polygonal description of the target geometry with a multipoint representation of dominant scattering centers, followed by image synthesis using the back-projection method. The sea surface is modeled as a distributed background that affects image contrast and structure; the study investigates how changes in range and azimuth resolution influence the visibility of structural features and the robustness of extracted characteristics in sea clutter. A MATLAB-based simulation framework is implemented, including reflected-signal generation, windowing, and full-polarimetric signature formation for a set of viewing angles. The results show that the multipoint model provides information content comparable to the polygonal model while substantially accelerating large-scale data generation and enabling the justification of resolution requirements for detection and classification of surface objects.

ПРИМЕНЕНИЕ МЕТОДА МНОГОСИГНАЛЬНОЙ КЛАССИФИКАЦИИ ДЛЯ ПЕЛЕНГОВАНИЯ ШУМОПОДОБНЫХ СИГНАЛОВ

асп. Комов С.Д.[1], к.т.н. Глухов А.Ю[1]

[1] ФГУП Всероссийский научный исследовательский институту автоматики им. Духова

В работе рассматривается применение метода многосигнальной классификации для пеленгования шумоподобных сигналов. Проведен ряд численных экспериментов при различных соотношениях длинны волны к диаметру антенной решетки, а также по воздействию на сигнал аддитивного шума и эффекта Доплера. Данные полученные в экспериментах показывают, что в сложных условиях распространения сигнала с значительным коэффициентом Доплера и низким отношением сигнал-шум, с использованием метода многосигнальной классификации, удается достаточно точно определить направление на источник сигнала.

Введение

Определение направления прихода сигнала (пеленгование) в многоканальных системах, построенных на основе разнесенных в пространстве антенн является одной из центральных задач наблюдения и охраны подводного и надводного пространства. В зависимости от различных условий применения используются активные и пассивные методы пеленгования. В работах [1],[2] подчеркивается перспективность пассивных методов в виду простоты их аппаратной реализации и возможности скрытного мониторинга.

В качестве источников информации при пассивном пеленговании могут использоваться широкополосные связные сигналы. Например, современные стандарты и коммерческие системы связи преимущественно используют шумоподобные сигналы [12], что требует применения специальных методов пространственной обработки [3],[4]. В литературе описаны различные методы пеленгования по задержкам сигнала [5], суммирования по методу максимального правдоподобия [6], корреляционно-интерферометрические методы пеленгования [7], а также спектральные методы на основе оценки пространственного спектра [8].

Существенное преимущество в условиях ограниченных физических размеров антенной решетки (АР) обеспечивают методы сверхразрешения, основанные на разделении подпространств сигнала и шума. К таким методам относится метод многосигнальной классификации MUSIC, широко представленный в работах [9-11 и др.]. Несмотря на наличие публикаций, посвящённых применению MUSIC к широкополосным сигналам, вопросы использования корреляционной информации, формируемой согласованными фильтрами (корреляционными приемниками) при приёме шумоподобных сигналов, остаются недостаточно раскрытыми.

Целью работы является разработка и исследование алгоритма пеленгования широкополосных шумоподобных сигналов на основе метода MUSIC с учётом особенностей корреляционной обработки.

Теоретические сведения

Шумоподобными сигналами (ШПС) называют класс широкополосных сигналов, для которых произведение эффективной ширины спектра F на длительность сигнала T значительно превышает единицу, то есть $FT \gg 1$ [4]. К основным свойствам и преимуществам шумоподобных сигналов относятся:

– высокая помехоустойчивость, обусловленная большой временно-частотной базой сигнала и возможностью применения корреляционных методов приёма;

– возможность одновременной работы нескольких абонентов в одном частотном диапазоне за счёт использования кодового разделения каналов;

– повышенная устойчивость к многолучевому распространению, обеспечиваемая возможностью временного и корреляционного разделения отражённых сигналов.

Благодаря совокупности перечисленных свойств шумоподобные сигналы получили широкое распространение в системах связи, навигации и радиолокации, функционирующих в условиях интенсивных помех и сложного распространения сигнала.

Прием такого сигнала осуществляется в несколько этапов:

– выделение окна анализа приемника;
– выделение комплексной огибающей;
– нахождение периодической взаимно-корреляционной функции (ПВКФ);
– принятие решения о приеме сигнала.

Для выделения комплексной огибающей используется гетеродин на нескольких частотных и временных каналах (f,τ), ФНЧ и механизм передискретизации [14].

Из-за корреляционных свойств М-последовательности ее ПВКФ имеет величину боковых пиков близкую к $1/\sqrt{N}$, где N – размер последовательности, а величину пика до 2^{Mpow}, где $Mpow$ – степень порождающего полинома, что позволяет достаточно точно определить сигнал по телу неопределенности, составленному из временных и частотных каналов (f,τ) с помощью решающего устройства [4].

Выходной информацией приемника является комплексная ПВКФ. Фаза ее пика соответствует разнице фаз пришедшего сигнала и гетеродина, а амплитуда пика содержит информацию об амплитуде сигнала.

Из-за сохранения информации о фазе и амплитуде ШПС в ПВКФ полученных корреляционными приемниками, для пеленгации такого сигнала целесообразно использовать метода MUSIC с использованием корреляционных приемников.

Рассмотрим антенную решетку, состоящую из M элементов. Для работы метода MUSIC минимальное расстояние между элементами должно быть меньше половины длины волны принимаемого сигнала. Сигнал от точечного источника падает на АР в виде плоского волнового фронта под углом θ. Сигналы с элементов АР можно рассматривать как:

$$\boldsymbol{X}(t) = a(f_c,\theta) * \boldsymbol{S}(t) + \boldsymbol{N}(t),$$

где $\boldsymbol{X}(t)$ – матрица наблюдаемых сигналов, $\boldsymbol{N}(t)$ – матрица аддитивного шума, $\boldsymbol{S}(t) = [s_1, s_2, \dots, s_n]^T$ – матрица сигналов источника, a – направляющий вектор, зависящей от несущей частоты сигнала f_c:

$$a(f_c,\theta) = [e^{-j2\pi f_c \Delta t_1(\theta)}, \dots, e^{-j2\pi f_c \Delta t_M(\theta)}] \quad (1)$$

Для фильтрации пришедших сигналов используется БПФ полосовой фильтр настроены на частоту гармонического сигнала.

Полная информация о направлении прихода сигнала содержится в корреляционной матрице сигнала:

$$\boldsymbol{R_{xx}} = \boldsymbol{Y} * \boldsymbol{Y}^H,$$

где $\boldsymbol{Y} = \hat{F}\{\boldsymbol{X}(t)\}$, $\hat{F}\{\}$ – оператор, преобразующий матрицу сигналов. В классическом алгоритме MUSIC $\hat{F}\{\}$ – оператор преобразования Фурье.

При разложении такой матрицы на собственные числа и вектора можно выделить подпространство шума и сигналов по значению собственных чисел матрицы. На основе ортогональности между подпространством шума и сканирующим направляющим векторам [13] сигналам строится функция:

$$P(\varphi, f_c) = \frac{1}{a(\varphi,f_c)^H \boldsymbol{E}_n \boldsymbol{E}_n^H\, a(\varphi,f_c)}, \quad (2)$$

где E_n – шумовое подпространство, $a(\varphi, f_c)$ – направляющий вектор, рассчитанный по формуле 1.

По максимуму функции $P(\varphi, f_c)$ определяется угол. Данный метод позволяет определить углы прихода плоского волнового фронта на АР если минимальное расстояние между элементами АР меньше половины длинны волны, что делает его удобным для построения систем пеленгования сигналов с антенными решетками с ультракороткой базой.

Описание компьютерной модели

На рисунке 1 приведена структурная схема алгоритма пеленгования, использующегося в компьютерной модели пеленгования ШПС на основе метода MUSIC.

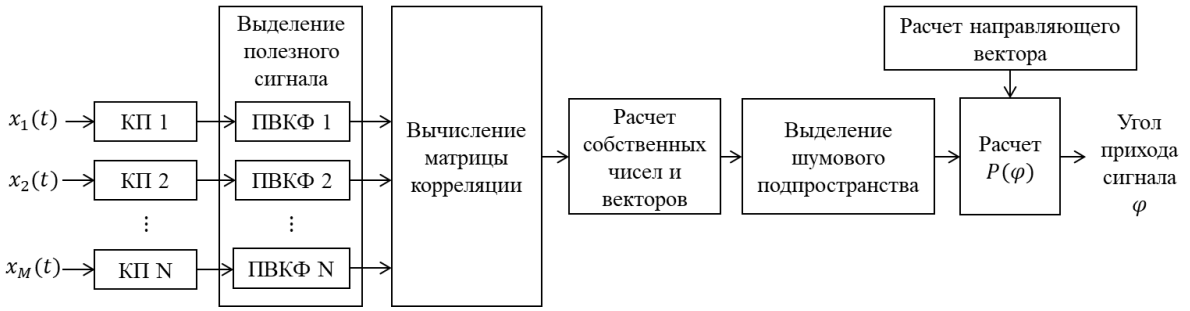


Рисунок 1. Структурная схема алгоритма MUSIC пеленгации ШПС.

Корреляционные приемники (КП) с синхронным гетеродином принимают сигналы с элементов АР $x_1(t), x_2(t) \dots x_M(t)$. После нахождения сигналов на всех приемниках по ПВКФ рассчитывается матрица корреляции $R_{xx}(\theta)$. Далее выделяются собственные числа λ и вектора $\bar{V}$ корреляционной матрицы. После нахождение собственных чисел и векторов определяется подпространства шума и сигнала для данной корреляционной матрицы. После расчет направляющих векторов на различных предполагаемых углах прихода сигнала $a(\varphi)$ и составлении функции (2) находится ее максимум. Аргумент максимума функции $P(\varphi)$ принимается за угол прихода сигнала.

На рисунке 2 представлена структурная схема формирователя ШПС сигналов с учетом эффектов среды (аддитивный белый шум и эффект Доплера).

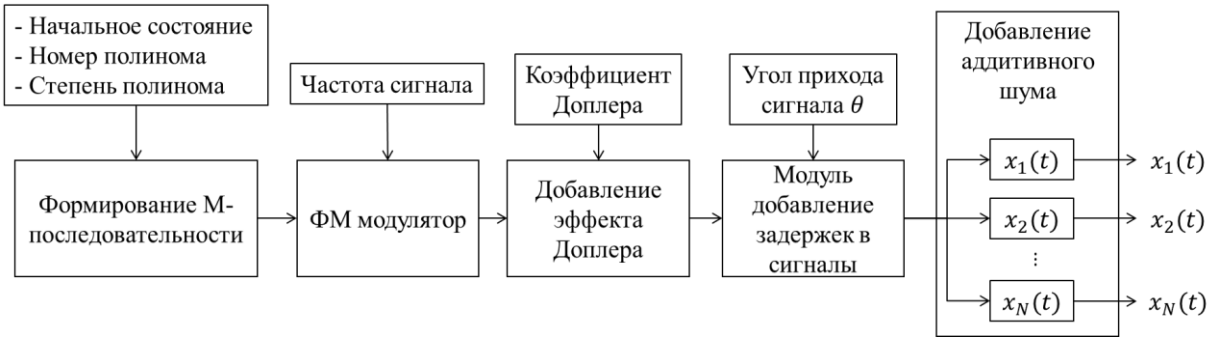


Рисунок 2. Структурная схема формирователя ШПС.

На ФМ модулятор поступает М последовательность с заданными начальным состоянием, номером и степенью полинома. После прохождения модулятора добавляется эффект Доплера с заданным коэффициентом Доплера $k_д = \frac{V}{c}$, где V – относительная скорость абонентов, c – скорость волнового фронта. Модуль добавления задержек сигнала для круговой АР формирует задержки по формуле:

$$dt_i = (-\cos(2\pi\theta) * x_i - \sin(2\pi\theta) * y_i) * \frac{1}{c},$$

где: θ – задаваемый угол прихода сигнала, x_i, y_i – координаты элемента АР, c – скорость волнового фронта.

Далее к сигналам добавляется аддитивный белый шум, нормированный на полосу сигнала и с заданным ОСШ.

Проведение численных экспериментов

Для проверки качества работы модели при воздействии аддитивного шума и эффекта Доплера на сигнал, а также при различных соотношениях диаметра АР к длине волны проведен ряд численных экспериментов.

Результаты экспериментов, проведенных с различным соотношением диаметра АР к длине волны сигнала показаны на рисунке 3. Сигналы формировались без шума и эффекта Доплера.

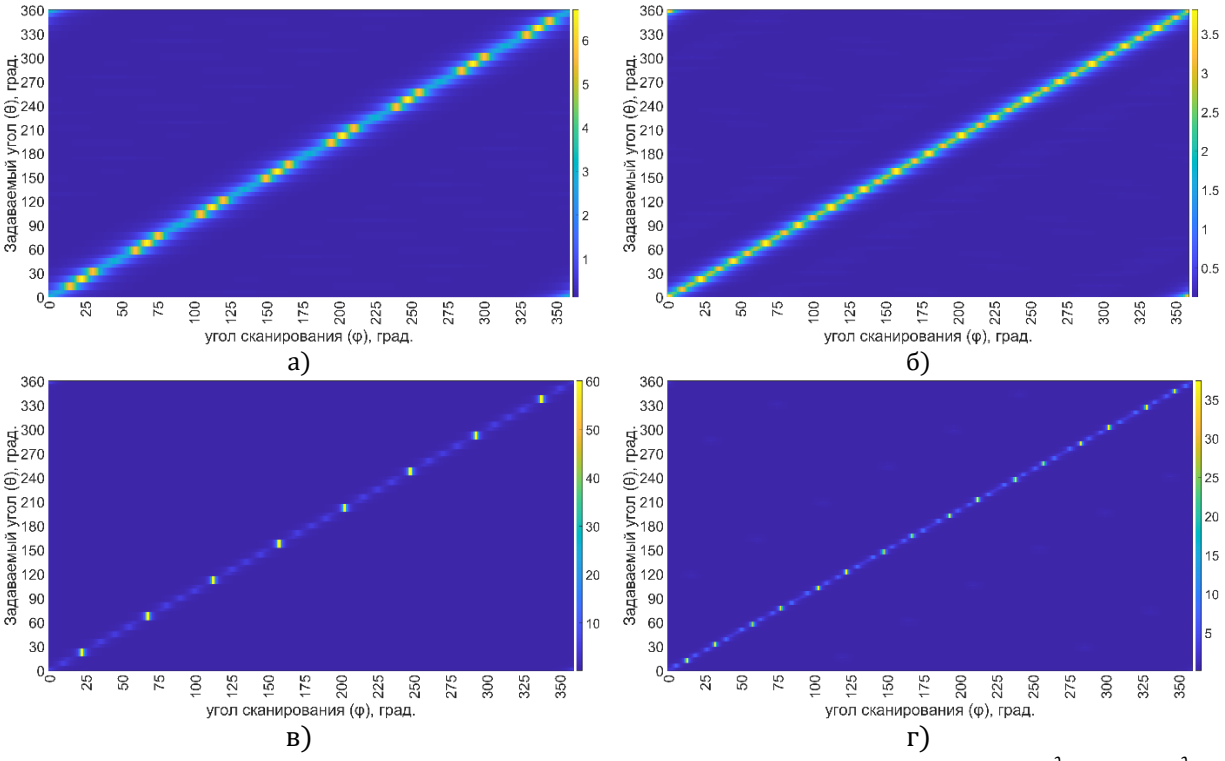


Рисунок 3. Зависимость функции $P(\varphi)$ от задаваемого угла θ при: а) $D = 0.7 * \frac{\lambda}{2}$, б) $D = \frac{\lambda}{2}$, в) $D = 1.14 * \frac{\lambda}{2}$, г) $D = \lambda$.

На рисунке изображена $P(\varphi)$, ось OX – углы сканирования пространства (φ), ось OY задаваемые углы прихода сигнала (θ). Из рисунка 3 видно, что при соотношении $D > \frac{\lambda}{2}$, где D – диаметр АР отношения пика функции $P(\varphi)$ к ее среднеквадратичному значению имеет зависимость от угла прихода сигнала. При соотношениях диметра антенной решетки равному половины длинны волны $D = \frac{\lambda}{2}$ данной зависимости не наблюдается. Зависимость отношения пика функции $P(\varphi)$ к ее среднеквадратичному значению для диаметра АР меньше половины длинны $D < \frac{\lambda}{2}$ связанно с низкой точностью формирования задержек при использовании данной КМ. Оптимальным решением является соотношение $D = \frac{\lambda}{2}$.

На рисунке 4. Приведен график зависимости усредненной ошибки определения угла от приведенного к полосе ОСШ сигнала.

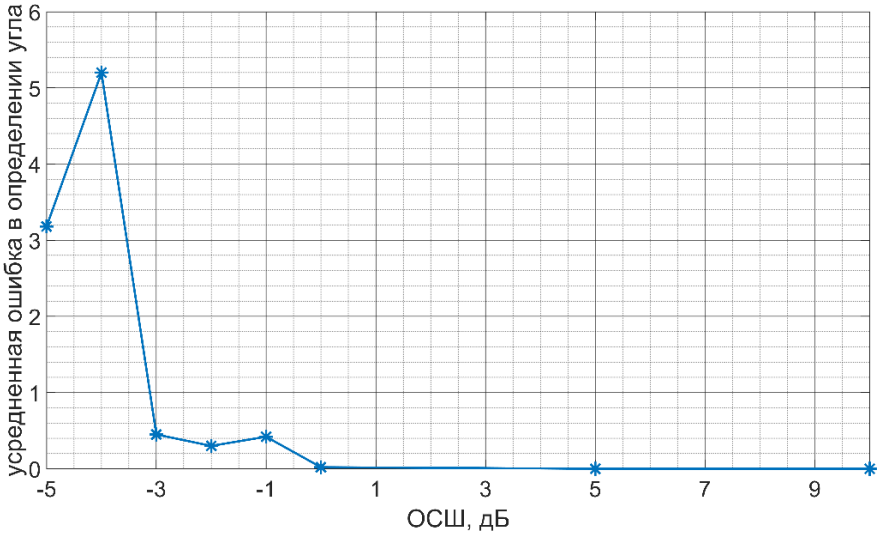


Рисунок 4. Зависимость ошибки определения угла от ОСШ сигнала.

Из графика видно, что усредненная абсолютная ошибка в определении угла при ОСШ 0 составляет 0.02 град, что позволяет точно определять углы прихода сигнала с приведенным ОСШ более 0 дБ.

Результаты эксперимента, проведенного при воздействии на сигнал аддитивного белого шума и эффекта Доплера приведены на рисунке 5.

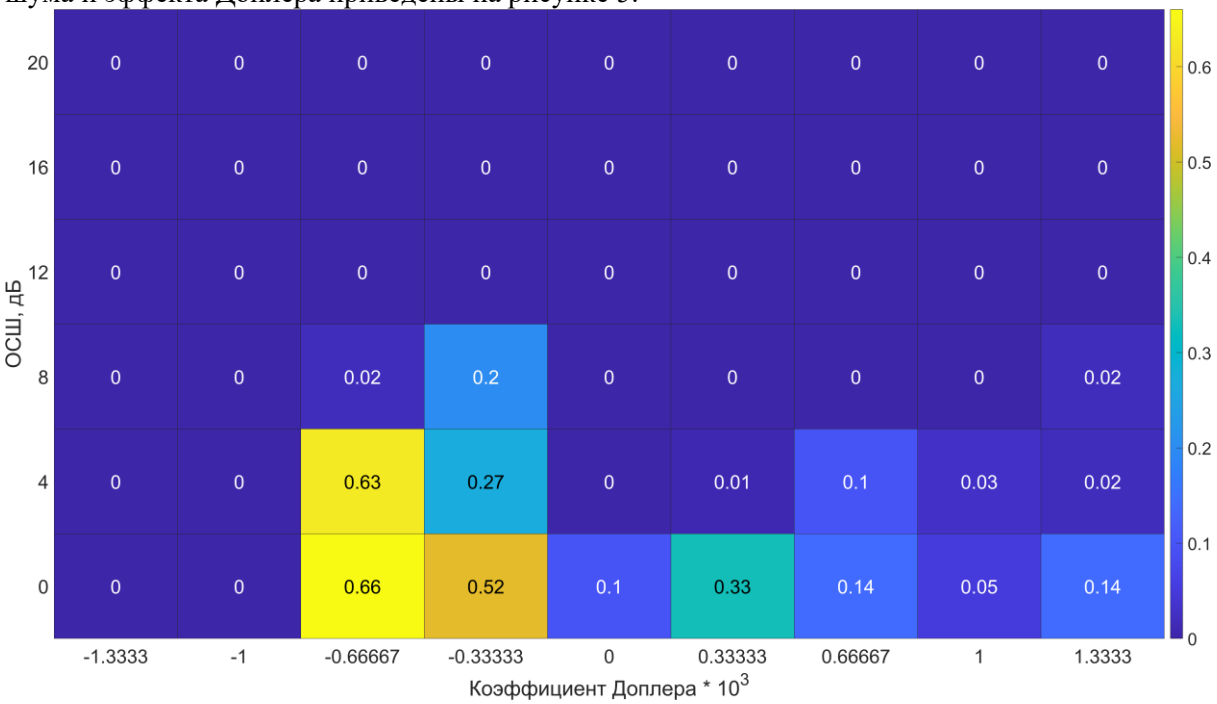


Рисунок 5. Зависимость ошибки определения угла от коэффициента Доплера и ОСШ сигнала.

Из рисунка видно, что зависимость определения угла от эффекта Доплера не значительная. Это связанно с учётом эффекта Доплера в корреляционных приемниках сигнала. Для уменьшения влияния коэффициента Доплера на метод MUSIC с использованием корреляционных приемников необходимо повышать количество доплеровских каналов при фиксированном Доплеровском коридоре.

Заключение

Основным отличием применения метода многосигнальной классификации применительно к ШПС является составление корреляционной матрицы не из спектров сигнала, а из ПВКФ корреляционных приемников. Для проверки качества работы метода был проведен ряд численных экспериментов результаты которых показывают, что данный метод пеленгации ШПС является устойчивым к аддитивным шумам и эффекту Доплера в системе, что может быть полезно в сложных условиях эксплуатации.

Литература

1. Васильев А. В., Колбаско И. В. Применение пассивных и активно-пассивных радиолокационных средств в зенитных ракетных войсках // Военная мысль. – 2023. – №. 6. – С. 73-79.

2. Петров Ю. В., Прошнов М. И. Пассивная пеленгация бпла с помощью sdr приёмника // Главный редактор Савочкин АА, канд. техн. наук, доцент, профессор кафедры «Радиоэлектроника и телекоммуникации» ФГАОУ ВО «Севастопольский государственный университет» Научный редактор Афонин ИЛ, д-р техн. наук, профессор, заведующий кафедрой «Радиоэлектроника и телекоммуникации» ФГАОУ ВО «Севастопольский государственный. – С. 27.

3. Гуляев Ю. В. и др. Информационные технологии на основе шумоподобных сигналов //Известия высших учебных заведений. Прикладная нелинейная динамика. – 2025. – Т. 33. – №. 5. – С. 629-656.

4. Варакин Л. Е. Системы связи с шумоподобными сигналами. – 1985.

5. Кучумов А. А., Мячин К. В., Припутин В. С. Аналитический метод пеленгации радиосигнала на равномерную круговую антенную решетку с произвольным количеством элементов //T-Comm-Телекоммуникации и Транспорт. – 2010. – №. 11. – С. 39-41.

6. Бобков Н. И., Габриэльян Д. Д., Иванов Ю. В. Амплитудно-фазовое распределение для формирования частотно-независимых диаграмм направленности линейных антенных решеток //Журнал радиоэлектроники. – 2019. – №. 2. – С. 10-10.

7. Ципоренко В. В. и др. Анализ точности беспоискового цифрового метода корреляционно-интерферометрического пеленгования с двухмерной корреляционной обработкой пространственного сигнала //Вісник Національного технічного університету України Київський політехнічний інститут. Серія: Радіотехніка. Радіоапаратобудування. – 2018. – №. 72. – С. 23-31.

8. Коптев В. А. Теоретические основы, области применения и характеристики методов сверхразрешения для измерительных радиотехнических систем //Экономика и качество систем связи. – 2025. – Т. 2. – №. 36. – С. 63-79.

9. Di Claudio E. D., Parisi R., Jacovitti G. Space time MUSIC: Consistent signal subspace estimation for wideband sensor arrays //IEEE Transactions on Signal Processing. – 2018. – Т. 66. – №. 10. – С. 2685-2699.

10. Xu W. et al. A Novel Wide-Band Directional MUSIC Algorithm Using the Strength Proportion //Sensors. – 2023. – Т. 23. – №. 9. – С. 4562.

11. Santosh S., Sahu O. P. Wideband Direction of Arrival (DOA) Estimation: A Comparative Study of Wideband MUSIC Method, Joint Diagonalization Structure (JDS) Method and Beam-Space Genetic Algorithm (BGA). – 2016.

12. 3GPP technical specifications define the standards for mobile telecommunications such as LTE and 5G NR [3GPP Specifications By Series, 2026].

13. Zhang X., Feng D. An efficient MUSIC algorithm enhanced by iteratively estimating signal subspace and its applications in spatial colored noise //Remote Sensing. – 2022. – Т. 14. – №. 17. – С. 4260.

14. Бобровников А. В., Силин А. С. Ускоренная синхронизация м-последовательностей //Вестник Воронежского государственного технического университета. – 2025. – Т. 21. – №. 2. – С. 207-211.

APPLICATION OF THE MULTI-SIGNAL CLASSIFICATION METHOD FOR TRACKING NOISE-LIKE SIGNALS

Komov S.D., Glukhov A.U.

FSUE Dukhov automatics research institute

The paper considers the application of the multi-signal classification method for tracking noise-like signals. A number of numerical experiments have been carried out with different ratios of wavelength to antenna array diameter, as well as on the effect of additive noise and the Doppler effect on the signal. The experimental data show that in difficult conditions of signal propagation with a significant Doppler coefficient and a low signal-to-noise ratio, using the multi-signal classification method, it is possible to accurately determine the direction to the signal source.

АЛГОРИТМ ВЫДЕЛЕНИЯ ПРИЗНАКОВ В СИГНАЛЕ СЕРДЦЕБИЕНИЯ ПРИ ИСПОЛЬЗОВАНИИ РАДИОЛОКАЦИОННОГО ДАТЧИКА

студ. Артемьев Д.Д.[1], доц. Витязев С.В.[1], Сысоев А.Н.[2]

[1] Рязанский государственный радиотехнический университет им. В.Ф. Уткина
[2] АО «Концерн «Автоматика»

Рассматривается алгоритм выделения компонент, соответствующих сердцебиению, в сигнале, регистрируемом с помощью радиолокационного датчика, с последующим формированием признаков, используемых в целях идентификации человека. Преимуществами данного способа регистрации сигнала и идентификации являются отсутствие необходимости в контакте датчиков с телом человека и невозможность подделки признаков, связанных с сердцебиением. Алгоритм построен на применении мел-фильтрации спектрограммы второй производной регистрируемого сигнала. Экспериментально показано, что данный подход может давать достоверные результаты идентификации.

Введение

В современном мире всё более распространенной является задача идентификации человека. На данный момент наиболее часто встречаются методы, основанные на использовании отпечатков пальцев, систем распознавания лиц или радужной оболочки глаз [1]. Недостатками этих подходов являются, соответственно, необходимость контакта датчика с телом человека, зависимость качества распознавания от освещенности, возможность фальсификации с помощью фотографии глаза. В последние годы распространение находят новые принципы идентификации. В частности, в качестве критериев могут использоваться особенности сердцебиения человека. Характер сердцебиения оказывается строго индивидуальным и устойчивым к попыткам фальсификации [2].

Зарегистрировать сердцебиение можно несколькими способами, базовым из которых является электрокардиография. Альтернативным способом является применение радиолокационного датчика [3, 4]. Такой способ регистрации имеет преимущества бесконтактной и при необходимости скрытной работы. В работе [5] предложен алгоритм идентификации на основе радиолокационных данных и показана высокая достижимая точность его работы.

В данной работе проводится разработка алгоритма обработки радиолокационного сигнала с целью выделения составляющих сердцебиения; разрабатывается стенд регистрации реальных данных; производится их обработка, и оценивается потенциальная возможность применения данного класса алгоритмов для решения задачи идентификации.

Описание алгоритма решения задачи

Последовательность этапов алгоритма обработки сигнала показана на Рисунке 1 [5].

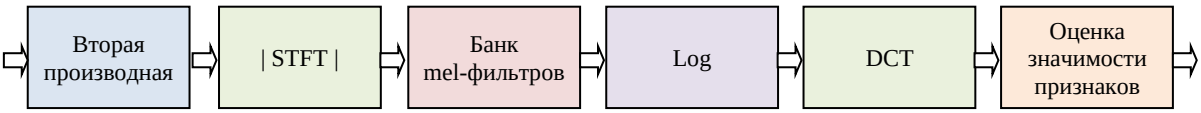


Рисунок 1 – Последовательность этапов алгоритма обработки сигналов

Входной сигнал представляет собой последовательность цифровых отсчетов, поступающих с заданной частотой дискретизации F_s в течение интервала наблюдения T_o. Компоненты сердцебиения в сигнале занимают более высокочастотный диапазон п сравнению с дыханием и другими мешающими факторами, поэтому имеет смысл перейти ко второй производной сигнала для усиления данных компонентов. Следующий этап алгоритма – это формирование спектрограммы (Short Time Fourier Transform – STFT) и взятие модуля, что соответствует спектральному анализу обрабатываемого сигнала с учетом динамики спектральных компонент. Спектрограмма строится с длиной окна L и перекрытием соседних окон на W процентов. Банк мел-фильтров – это набор полосовых фильтров с неравномерно распределенными по оси частот полосами пропускания и коэффициентами усиления, позволяющий сжать спектр сигнала до

ограниченного набора спектральных коэффициентов с регулируемыми весами: небольшой набор коэффициентов отражает весь спектральный состав сигнала с учетом важности тех или иных субполос, а также с усреднением результатов по временным интервалам, для которых строится спектрограмма. Здесь следует сделать замечание, что традиционно мел-фильтрация применяется при анализе спектральных характеристик речевых сигналов. Однако, как указано в [5], сигналы сердцебиения, зарегистрированные радиолокационным датчиком, имеют схожую с речевыми сигналами структуру периодических на коротких временных интервалах колебаний, и применение мел-шкалы для таких сигналов оказывается целесообразным. Последними этапами обработки являются взятие десятичного логарифма (Log) спектральных коэффициентов и дискретное косинусное преобразование (DCT). В целом данный алгоритм соответствует кепстральному анализу, нашедшему успешное применение в задачах обработки речевых сигналов [6]. Считается, что обработка радиолокационных сигналов сердцебиения может быть схожа с данными подходами. После формирования признаков указанным способом остается провести оценку их значимости.

Экспериментальный стенд и результаты исследования

Для анализа эффективности описанного алгоритма был построен экспериментальный стенд на базе отладочного модуля обработки радиолокационных сигналов с видом модуляции FMCW, рабочей частотой 77 ГГц, шириной полосы до 4 ГГц, разрешением по дальности до 4 см, восемью принимающими антеннами в режиме MIMO и комплексным представлением сигналов на выходе аналого-цифрового преобразователя [7]. С помощью данного модуля была произведена регистрация сигналов, отраженных от грудной клетки, для трех человек – Рисунок 2.

Сигналы регистрировались в одном канале дальности на расстоянии 20 см. Интервал наблюдения для каждого сигнала – 60 секунд. Частота дискретизации 125 Гц. Вид входного сигнала (фрагмент) представлен на Рисунке 3. Обработка производилась в системе моделирования на персональном компьютере, то есть не в реальном времени.

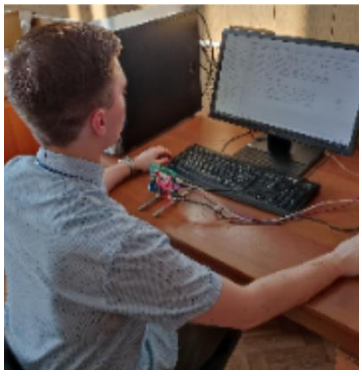


Рисунок 2 – Радиолокационный датчик и запись сигналов

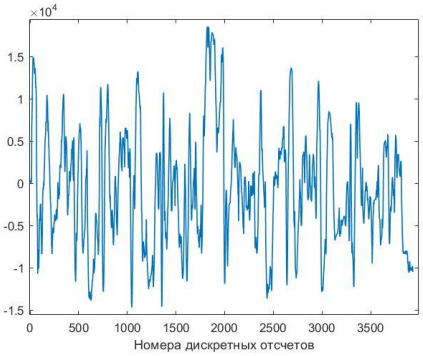


Рисунок 3 – Сигнал на входе

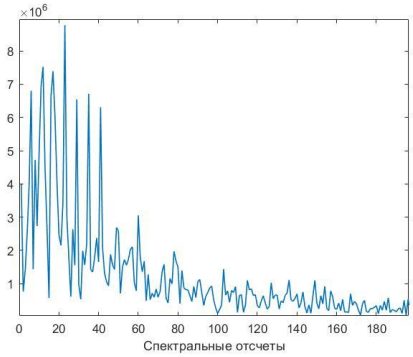


Рисунок 4 – Спектр входного сигнала

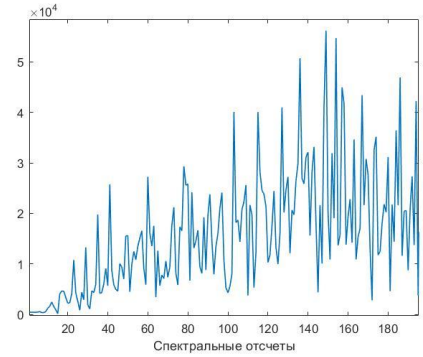


Рисунок 5 – Спектр второй производной сигнала

На Рисунке 4 показан спектр входного сигнала (фрагмент). Основными гармониками являются гармоники компонент дыхания и низкочастотные помехи. Усиление компонент, соответствующих сердцебиению, путем взятия второй производной иллюстрирует амплитудный спектр на Рисунке 5 (фрагмент спектра). Спектрограмма получаемого таким образом сигнала показана на Рисунке 6. Выход банка мел-фильтров проиллюстрирован на Рисунке 7.

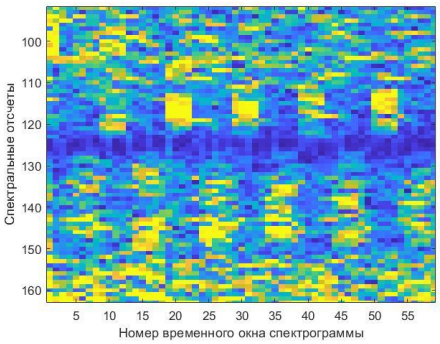


Рисунок 6 – Спектрограмма

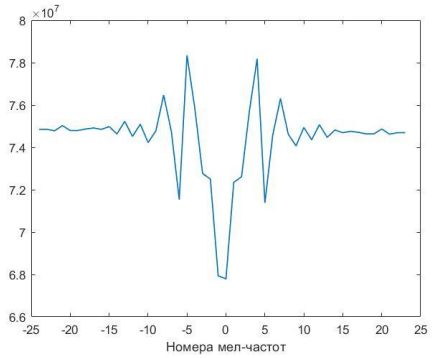


Рисунок 7 – Выход банка мел-фильтров

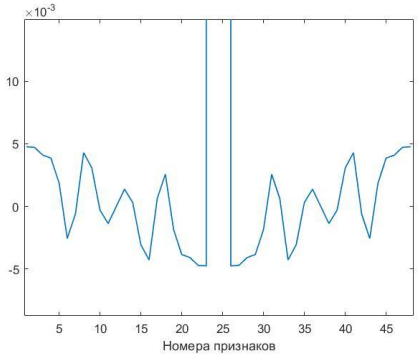


Рисунок 8 – Признаки для идентификации

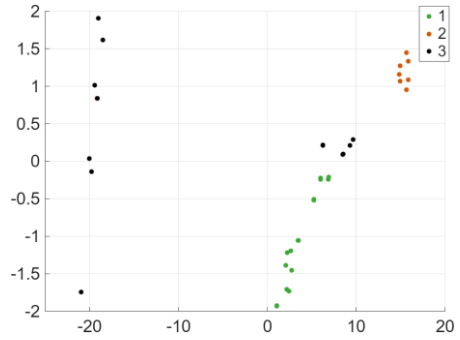


Рисунок 9 – Классификация процедурой SNE

На выходе банка мел-фильтров основной вес имеют те компоненты, которые соответствуют сердцебиению. На Рисунке 8 показан результат обработки – вектор коэффициентов, используемых в качестве признаков для идентификации.

Сама идентификация в рамках данной работы не проводится. Для анализа эффективности исследуемого алгоритма сформированные наборы данных для тестирования подавались на процедуру SNE (Stochastic Neighbor Embedding), оценивающую и визуализирующую различимость векторов признаков, относящихся к разным объектам идентификации (Рисунок 9). В данном случае записи выполнялись для трех разных объектов – трех человек. Для каждого человека было сделано 16 записей сигнала. По рисунку видно, что векторы признаков в основном правильно разделяются в три разные группы, хотя присутствуют и перекрытия разных классов. Можно сделать вывод о работоспособности предлагаемого алгоритма, но, в то же время, о необходимости проведения дополнительных исследований и повышения эффективности рассматриваемого алгоритма.

Заключение

Таким образом, в работе описана разработка алгоритма выделения признаков из сигнала сердцебиения, регистрируемого радиолокационным датчиком, в целях решения задачи идентификации человека. Приведено описание алгоритма обработки сигнала. Создан стенд и произведена регистрация реальных данных для трех человек. Выполнена обработка записанных сигналов в соответствии с представленным алгоритмом. Сделан вывод о потенциальной работоспособности предлагаемого решения и, одновременно, о необходимости проведения дополнительных исследований и повышения эффективности обработки сигналов.

Литература

1. Кухарев, Г. А. Методы обработки и распознавания изображений лиц в задачах биометрии / Г. А. Кухарев, Е. И. Каменская, Ю. Н. Матвеев, Н. Л. Щеголева ; под ред. М. В. Хитрова. — Санкт-Петербург : Политехника, 2013. — 388 с. : ил.
2. L. Biel, O. Pettersson, L. Philipson and P. Wide, "ECG analysis: a new approach in human identification," in IEEE Transactions on Instrumentation and Measurement, vol. 50, no. 3, pp. 808-812, June 2001, doi: 10.1109/19.930458.
3. Q. He, E. Yang, and S. Fang, ‘‘A survey on human profile information inference via wireless signals,’’ IEEE Commun. Surveys Tuts., vol. 26, no. 4, pp. 2577–2610, 4th Quart., 2024, doi: 10.1109/COMST.2024.
4. G. Paterniani, D. Sgreccia, A. Davoli, G. Guerzoni, P. Di Viesti, A. C. Valenti, M. Vitolo, G. M. Vitetta, and G. Boriani, ‘‘Radar-based monitoring of vital signs: A tutorial overview,’’ Proc. IEEE, vol. 111, no. 3, pp. 277–317, Mar. 2023, doi: 10.1109/JPROC.2023.3244362.
5. H. Kobayashi, Y. Tanaka and T. Sakamoto, "Individual Identification Using Radar-Measured Respiratory and Heartbeat Features," in IEEE Access, vol. 12, pp. 190972-190987, 2024, doi: 10.1109/ACCESS.2024.3517668.
6. Z. K. Abdul and A. K. Al-Talabani, "Mel Frequency Cepstral Coefficient and its Applications: A Review," in IEEE Access, vol. 10, pp. 122136-122158, 2022, doi: 10.1109/ACCESS.2022.3223444.
7. Texas Instruments, "AWR1642 Evaluation Module (AWR1642BOOST) Single-Chip mmWave Sensing Solution User's Guide," SWRU508B, May 2017, Revised April 2018.

AN ALGORITHM FOR FEATURE EXTRACTION IN A HEARTBEAT SIGNAL WHEN USING A RADAR SENSOR

stud. Artemyev D.D.[1], assoc. prof. Vityazev S.V.[1], Sysoev A.N.[2]

[1] Ryazan State Radio Engineering University named after V.F. Utkin
[2] JSC Concern Avtomatika

An algorithm for isolating the components of a heartbeat in a signal received using a radar sensor and forming signs used to identify a person is considered. The advantages of this method of signal registration and identification are the absence of the need for sensors to contact the human body and the inability to fake signs associated with a heartbeat. The algorithm is based on the use of chalk filtering of the spectrogram of the second derivative of the recorded signal. It has been experimentally shown that this approach can give reliable identification results.

СИСТЕМНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ СРЕДСТВ ЭКВИВАЛЕНТНЫХ КОДОВ

проф. Гладких А.А., асс. Брынза А.А., асс. Бакурова А.Д.

Ульяновский государственный технический университет

Для реализации высоких темпов обработки цифровых данных в современных телекоммуникационных системах становятся востребованы методы обработки данных с низкой задержкой (25 мкс ~ 1 мс) и одновременной актуализации требований по их достоверности. Подобные методы могут быть реализованы за счет уменьшения длины кодовых векторов избыточных кодов, но с обеспечением их достаточно высокой исправляющей способности. Одним из научных направлений поиска приемлемых технических решений в данной предметной области являются исследования, связанные с изучением возможностей средств перестановочного декодирования, основанного на теории получения для некоторого систематического блокового кода ограниченного подмножества эквивалентных ему кодов. Для формирований указанного подмножества используются, как правило, мягкие решения, применение которых по асимптотическим оценкам обеспечивает повышение энергетического выигрыша кода до 3 дБ. При этом в процедуре декодирования эквивалентного кода метрика Хэмминга заменяется более высоким параметром, равным числу введенных в код избыточных разрядов. Недостатком указанного метода обработки цифровых данных является существование второго замкнутого подмножества таких перестановок символов двоичных кодовых векторов, которые приводят к вырожденным порождающим матрицам вновь образованным кодам, соответствующих перестановкам наиболее надежным символам. И особенности именно таких перестановок в системе кодеков для различных кодов могут быть изучены только методами математической статистики для исследования сложных объектов как целостных систем.

Особенности статистики перестановок коротких блоковых кодов

Математические основы формирования эквивалентных кодов (ЭК) были заложены в работах [1, 2]. В них было показано, что подобные конструкции способны (при определенных условиях) превзойти возможности избыточных блоковых систематических кодов по исправлению ошибок за пределами метрики Хэмминга, основанной на понятии комбинаций минимального веса в весовом спектре конкретного кода. Если комбинации минимального веса имеют значение указанного параметра равное d_{min}, число гарантированно исправляемых кодом ошибок равно значению $t = \lfloor (d_{min} - 1)/2 \rfloor$. Было показано, что применение в системе защиты цифровых данных от ошибок при использовании метода перестановочного декодирования (ПД) обеспечивает исправление $(n - k)$ ошибок, что однозначно превосходит параметр t и отвечает требованиям систем сверхнадежной связи с низкой задержкой (URLLC, Ultra-Reliable Low-latency Communications) [3, 4]. Здесь n представляет длину кодового вектора, а k – число информационных элементов в нем. Развитие метода ЭК имеет первостепенное значение для многих предметных областей на фоне становления в них возможностей машинного обучения и средств тактильного интернета. Следует указать, что развитию метода ПД во многом препятствовали обязательные матричные вычисления, которые требовали большого вычислительного ресурса, предоставление которого, например, в системе беспилотных средств представляет самостоятельную специфическую задачу. По мете развития этого направления было доказано, что процедура получения множества ЭК может быть организовано заранее. Для этого для каждой перестановки нумераторов элементов кодовой комбинации избыточного кода априори определяется порождающая матрица соответствующего ей ЭК, содержание которой записывается в специальною карту памяти. Поскольку полученные матрицы должны быть представлены в систематической форме, то в карту памяти записывается только проверочная часть такой матрицы без учета единичной матрицы. Это способствует рациональному использованию памяти применяемого микропроцессора. Данные в такой карте могут по мере эксплуатации системы пополняться дополнительными сведениями поэтому название таких карт в работе сохранилось в виде «когнитивных карт», хотя, следуя терминологии предметной области когнитивных автоматов в случае исчерпывающей полноты данных такой карты, уместно было бы назвать такой элемент декодера обычной картой памяти. Двоичные коды не являются максимально декодируемыми, поэтому возможны положительные (производительные) и отрицательные (непроизводительные) решения по возможности создания ЭК. При этом следует учитывать, что для производительных перестановок переставленные информационные часть порождающих матриц не являются вырожденными и обеспечивают получение ЭК. Напротив,

если подобные матрицы оказываются вырожденными, то перестановка не обеспечивает получение требуемого кода. Такая перестановка нуждается в коррекции. Результаты статистических испытаний по перебору перестановок показаны на рисунке 1, где для каждого кода слева указано процентное содержание производительных перестановок нумераторов (ППН), а справа – непроизводительных перестановок нумераторов (НПН).



Рисунок 1 – Диаграммы процентных соотношений ППН и НПН для различных блоковых кодов

Таким образом, все множество перестановок P любого блокового кода разбивается на два непересекающихся подмножеств ППН $P_{\text{ППН}} \in P$ и $P_{\text{НПН}} \in P$. При этом $P_{\text{НПН}}$ может содержать в себе ряд подмножеств $P_{\text{НПН}} \subseteq P_1$ перестановки, требующие при своей коррекции только одного символа (как это происходит для всех НПН кода Хэмминга), но все другие коды потребуют определенный процент замен, при этом возможны варианты, когда необходимо будет искать вариант с двумя или пятью заменами. Все зависит от структуры проверочной части порождающей матрицы систематического блокового кода [5]. Последнее обстоятельство подчеркивает преимущество кода (7, 4, 3). При этом предлагается все операции по поиску группы надежных символов принятого кодового вектора осуществлять на основе текущих и принятых приемником данных, присущих именно данному кодовому вектору, который находится в обработке. Принципиально при такой постановке задачи, оперируя малым количеством мягких решений символов (МРС), возможно с определенной осторожностью использование параметров среднего арифметического, размах, медиана и мода, но они оказываются полезным в технологии ПД. Приведенные диаграммы показывают, что для указанных выше ограничений применение кодов (7, 4, 3) и (15, 5, 7) предпочтительно. Ниже приводятся данные, полученные в ходе моделирования гауссовского канала связи при использовании в системе ПД кода Хэмминга. Наибольший интерес вызывает процесс изменения распределения значений мягких решений в зависимости от отношения сигнал/шум. Результаты работы имитационной модели показаны в таблице.

Таблица. Статистика МРС на длине кодового вектора как функция отношения сигнал/шум.

Отношение сигнал/шум (дБ)	Среднее значение	Вероятность ошибки на бит	Частота МРС в %							
			7	6	5	4	3	2	1	0
10,0	6,8	4×10^{-6}	29,8	70,2	-	-	-	-	-	-
6,9	6,7	$7{,}8\times 10^{-4}$	7,5	92,4	0,1	-	-	-	-	-
5,2	6,5	5×10^{-3}	3,9	93,1	3,0	-	-	-	-	-
4,0	6,4	1×10^{-2}	2,9	85,2	11,9	0,1	-	-	-	-
3,0	6,2	2×10^{-2}	2,2	73,3	23,9	0,6	-	-	-	-
2,2	6,1	$3{,}4\times 10^{-2}$	1,9	61,5	34,2	2,4	-	-	-	-
1,5	5,9	$4{,}5\times 10^{-2}$	1,7	51,4	41,3	5,4	0,2	-	-	-
1,0	5,8	$5{,}7\times 10^{-2}$	1,5	43,1	45,2	9,4	0,6	-	-	-
0,5	5,7	$6{,}8\times 10^{-2}$	1,4	37,1	46,7	13,5	1,2	0,1	-	-
0,0	5,6	$7{,}8\times 10^{-2}$	1,3	32,3	47,1	17,2	2,1	0,1	-	-

В ходе испытаний оценивались данные для 10^6 комбинаций, статистика представительна.

Замечательным свойством предложенного подхода на базе среднего арифметического поиска перестановки вида ППН является исключение из алгоритма процедуры ранжирования нумераторов по значениям МРС, что приводит к экономии и координации вычислительного ресурса. При этом средний показатель может быть индикатором применения против системы обмена данными систем помех деструктивного характера с последующим изменением режима работы первой [7].

Управление группой робототехнических устройств на основе эквивалентных кодов

Принципа двойственности (суперсистемности) указывает на то, что объект исследования должен рассматриваться и как система, и как подсистема более высокого уровня надсистемы или суперсистемы. Известно, что перестановка размерности в k символов способна сформировать $k!$ подстановок. Обычно n – длина вектора избыточного кода превосходит значение k и если продвигать некоторую комбинация нумераторов, состоящую из k элементов циклически по номерам комбинации из n элементов, то однозначно может быть получено только n перестановок. Если исходная перестановка относилась к подмножеству $P_{\text{ППН}}$, то все перестановки из цикла так же будут принадлежать подмножеству $P_{\text{ППН}}$, но с другой структурой проверочных матриц на каждом шаге цикла [8].

Это правило дает возможность не применять систему ПД и поиска ЭК для каждой принятой комбинации, а создать систему управления, которая может быть классифицирована как целостная, многоуровневая сё возможностями гибкой динамической структуризации, исходя из требований целевой функции управления. Это подчеркивает совершенно новые эмерджентные свойства ЭК.

Принцип целенаправленной передачи команд управления отдельным элементам или кластерам в группе робототехнических объектов может использоваться при проектиров ании новых и модернизации существующих систем обмена данными в системах управления, например, в системах дистанционного контроля и управления робототехническими объектами или в системах управления самоорганизующейся группы робототехнических объектов. Реализация основана на использовании эквивалентных групповых систематических кодов, что существенно расширяет их возможности относительно систем ПД. Важно, что на основе основного блокового избыточного систематического кода возможно создание нескольких различных ЭК, которые могут быть использованы для управления группой объектов. Отличие этих кодов друг от друга выражается структурой их проверочных матриц, которая зависит только от системы перестановок срок, входящих в состав проверочной матрицы основного кода. Следовательно, по мере перехода от одной перестановки k нумераторов к другой в проверочной матрице кода (7, 4, 3) (по крайней мере это установлено точно меняются местами только строки исходной проверочной матрицы основного корда). Пример структуры кластера кода (7, 4, 3) показан в таблице 2.

Таблица 2. Пример структуры одного из кластеров кода (7, 4, 3)

Значения перестановок нумераторов кластера S_i						
1246	**2356**	**3467**	**1457**	**1256**	**2367**	**1347**
Содержание соответствующих проверочных матриц эквивалентных кодов						
$\begin{matrix}0&1&1\\(1&0&1)\\1&1&1\\1&1&0\end{matrix}$	$\begin{matrix}1&0&1\\(0&1&1)\\1&1&1\\1&1&0\end{matrix}$	$\begin{matrix}1&1&0\\(1&0&1)\\0&1&1\\1&1&1\end{matrix}$	$\begin{matrix}0&1&1\\(1&1&0)\\1&0&1\\1&1&1\end{matrix}$	$\begin{matrix}1&0&1\\(0&1&1)\\1&1&0\\1&1&1\end{matrix}$	$\begin{matrix}1&1&1\\(1&0&1)\\0&1&1\\1&1&0\end{matrix}$	$\begin{matrix}1&1&0\\(1&1&1)\\1&0&1\\0&1&1\end{matrix}$

Это позволяет точно идентифицировать каждый объект или предписать такие матрицы группе объектов для их совместного действия в ходе решения задач целевой функции управления. Для рационального решения указанной задачи показано, что группу объектов за счет использования ЭК допускается разбивать на кластеры и связывать отдельные кластеры с друг с другом или осуществлять связь разных объектов, входящих в разные кластеры по принципу каждый с

каждым. Показывается возможность реализации таких соединений за счет использования уникальной структуры проверочных матриц ЭК. Следовательно, конфигурация строк проверочной матрицы из состава порождающей матрицы на своем передатчике является аналогом адреса вызываемого корреспондента. Следствием таких действий будет являться то, что только у корреспондента **1347** синдромы комбинаций избыточного кода будут равны нулевому значению. У других участников данного кластера этого не произойдет.

Заключение

Таким образом, с точки зрения системного анализа и их практического использования средства ЭК целесообразно рассматривать в узком и широком смысле со всеми их достоинствами и недостатками.

Литература

1. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. // М. : ИД Вильямс, 2003. 1104 с.
2. А.А. Гладких Перестановочное декодирование как инструмент повышения энергетической эффективности систем обмена данными // Электросвязь. 2017. № 8. С. 52–56.
3. А.Л.Х. Аттаби, А.А. Брынза, Д.В. Ганин, А.А. Ничунаев, А.В. Новоселов. Оценка статистических характеристик перестановочного декодера методом его программной реализации // Автоматизация процессов управления. 2023. № 2 (72). С. 91–98.
4. Д.В. Мишин, А.А. Гладких, В.И. Кутузов, А.Л.Х. Аттаби. Исследование процессов когнитивной обработки данных в системах радиосвязи с перестановочным декодированием // Физика волновых процессов и радиотехнические системы. 2024. Т. 27, № 1. С. 103–112.
5. А.С.А. Аль-Месри, Н.А. Пчелин, В.П. Смолеха, А.Д. Бакурова, С.В. Шахтанов. Экспериментальная оценка множества непроизводительных перестановок в процедуре поиска эквивалентных кодов // Автоматизация процессов управления. 2021. № 4 (66). С. 72-77.
6. А.Д. Бакурова. Расчет количества арифметических операций для поиска обратной матрицы и повышения надежности принятых символов в алгоритме перестановочного декодирования на примере кода Хэмминга (7,4,3) // Инфокоммуникационные технологии, 2024. Т.22, №4 (88), С. 42-47.
7. А.А. Брынза Априорное формирование когнитивной карты непроизводительных решений в системе перестановочного декодирования// Автоматизация процессов управления. 2024. № 4 (78). С. 18–23.
8. А.А. Гладких, А.А. Овинников, Н.А. Пчелин, А.А. Брынза. Перестановочное декодирование с системой адаптированных альтернативных решений // Цифровая обработка сигналов. 2023. № 4. С. 73–78.

SYSTEMATIC STATISTICAL ANALYSIS OF EQUIVALENT CODES

Gladkikh A.A., Brynza A.A., Bakurova A.D.

Ulyanovsk State Technical University

In order to realize high rates of digital data processing in modern telecommunication systems, low-latency data processing methods (25 microseconds ~ 1 ms) and simultaneous updating of requirements for their reliability are becoming in demand. Such methods can be implemented by reducing the length of the redundant code vectors, but ensuring their sufficiently high correcting power. One of the scientific directions of the search for acceptable technical solutions in this subject area is research related to the study of the possibili ties of permutation decoding based on the theory of obtaining a limited subset of equivalent codes for a certain systematic block code. For the formation of this subset, soft solutions are usually used, the use of which, according to asymptotic estimates, provides an increase in the energy gain of the code to 3 dB. In this case, in the procedure for decoding the equivalent code, the Hamming metric is replaced by a higher parameter equal to the number of excess digits entered into the code. The disadvantage of this method of processing digital data is the existence of a second closed subset of such permutations of binary code vector symbols, which lead to degenerate generating matrices of newly formed codes corresponding to permutations of the most reliable symbols. And the features of precisely such permutations in the codec system for various codes can only be studied using mathematical statistics methods to study complex objects as integral systems.

НЕЙРОСЕТЕВОЙ ТУРБО-ЭКВАЛАЙЗЕР С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОГО ОБУЧЕНИЯ НА ОСНОВЕ МЕХАНИЗМ SELF-ATTENTION

к.т.н. Малютин А.А.[1], Портниченко М.Д.[1], А.Н. Зюзин А.Н.[1], к.т.н., доц. Хорошайлова М.В.[2]

[1] АО НВП «ПРОТЕК», г. Воронеж, Россия
[2]Воронежский государственный технический университет, г. Воронеж, Россия

Предлагается турбо-эквалайзер с формированием мягких решений на базе нейросети Transformer-Lite, интегрированной в итеративный цикл сферического декодирования со стратегией перечисления Шнорра-Эйхнера в конфигурации K-best. Нейросеть использует механизм Self-Attention для учёта наличия станционных помех реального эфира. Обучение производится на гибридном датасете, включающем как синтетические данные, так и натурные записи с эфира. Алгоритм имеет полиномиальную и строго фиксированную вычислительную сложность.

Введение

Декаметровый (ДКМВ) диапазон радиосвязи остаётся критически важным для обеспечения различных видов связи, однако передача данных в нём сопряжена с серьёзными физическими ограничениями: многолучевым распространением сигнала и наличием интенсивных сосредоточенных по частоте станционных помех при относительно высоком уровне фонового шума. Современные требования к пропускной способности диктуют необходимость использования многопозиционных частотно-эффективных видов модуляции, требующих когерентного приёма. В этих условиях классические линейные эквалайзеры демонстрируют неудовлетворительную помехоустойчивость из-за эффекта усиления шума в провалах частотной характеристики. С другой стороны, оптимальное детектирование по критерию максимального правдоподобия (ML) (алгоритм Витерби или BCJR-MAP) обладает неприемлемо высокой вычислительной сложностью из-за большой памяти канала при требуемой кратности модуляции. Итеративные методы приёма хорошо решают проблему повышения помехоустойчивости и пропускной способности при приемлемой сложности реализации, но только при наличии помех, близких к белому гауссовскому шуму. Попытки разработать оптимальные алгоритмы приёма при наличии станционных помех хорошо известны, но приводят к сложно реализуемым алгоритмам, каждый из которых эффективен только при определённом виде помех, и вследствие этого вряд ли могут быть применены на практике.

Поскольку итеративные методы приёма до сих пор являются наиболее эффективным способом обработки при допустимой сложности реализующих их алгоритмов, задача их дальнейшего совершенствования в направлении лучшего учёта реальной помеховой обстановки является актуальной. Известно, что узким местом при использовании классических турбо-эквалайзеров последовательных модемов в реальных многолучевых каналах связи со станционными помехами является метод расчёта логарифмических отношений правдоподобия (LLR), который практически всегда выполняется в предположении гауссовской статистики помех и часто приводит к неверным решениям при отклонении статистики от изначально предполагаемой. А наличие именно мягких решений, подаваемых на вход декодера, является ключевым требованием, необходимым для достижения высоких характеристик помехоустойчивости. В недавнее время появилась возможность использования методов искусственного интеллекта, особенно методов глубокого обучения, для решения задач подобного рода [1]. В частности, известны методы расчёта LLR с помощью свёрточной нейросети глубокого обучения при наличии помех с негауссовским распределением. В данной работе предлагается более совершенный подход к расчёту LLR, адаптированный для использования именно в современном турбо-эквалайзере последовательного модема, построенном с использованием в качестве первой ступени схемы сферического декодирования (вместо линейного трансверсального фильтра классического турбо-эквалайзера) в сочетании с алгоритмом приведения базиса решётки с помощью процедуры Ленстра-Ленстра-Ловаса (LLL). Собственно, метод расчёта мягких решений состоит в использовании нейросети трансформерного типа, использующей механизм самовнимания (Self-Attention). Для слепой/полуслепой оценки параметров канала, уточняемой в ходе итераций, используется EM-алгоритм. В качестве базовой стратегии перебора узлов решетки принята стратегия Шнорра-Эйхнера, обеспечивающая упорядочивание кандидатов по возрастанию метрики расстояния

внутри каждого уровня поиска в начале цикла перебора. Для обеспечения полиномиальной сложности и строго фиксированного времени вычислений алгоритм реализован в конфигурации K-best.

Задача детектирования как поиск ближайшего вектора в решетке

В системах связи по каналам с межсимвольной интерференцией принятый сигнал описывается вектором $\mathbf{y} \in \mathbb{C}^M$, который представляет собой суперпозицию передаваемых символов, искаженных ISI каналом с матрицей $\mathbf{H}$. Задача оптимального детектирования по критерию максимального правдоподобия формулируется как минимизация евклидова расстояния: $\hat{\mathbf{x}}_{ML} = \arg\min_{\mathbf{x}\in\mathcal{S}^N} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$ где $\mathcal{S}^N$ — дискретное множество векторов, элементами которых являются точки созвездия. Матрицу канала $\mathbf{H}$ можно рассматривать как базис некоторой N-мерной решетки $\mathcal{L}$ в комплексном пространстве, которая определяется как совокупность всех возможных линейных комбинаций столбцов матрицы $\mathbf{H}$ с целыми или дискретными комплексными коэффициентами: $\mathcal{L}(\mathbf{H}) = \left\{\mathbf{H}\mathbf{x} \mid \mathbf{x} \in \mathcal{S}^N\right\}$. Таким образом, каждый переданный вектор $\mathbf{x}$ отображается в определенную точку решетки $\mathbf{v} = \mathbf{H}\mathbf{x}$ в пространстве сигнала. Принятый сигнал $\mathbf{y}$ является точкой этого пространства, смещенной относительно узла решетки на вектор шума $\mathbf{n}$. С математической точки зрения задача идентична классической задаче поиска ближайшего вектора (Closest Vector Problem, CVP) в решетке. Узлы решетки — это все возможные варианты сигналов, которые могли быть получены приемником при отсутствии шума. Целевой вектор — это фактически принятый сигнал $\mathbf{y}$. Сложность решения задачи CVP напрямую зависит от ортогональности базиса $\mathbf{H}$. Из-за многолучевости столбцы матрицы $\mathbf{H}$ могут быть сильно коррелированы, что вызывает «плохую обусловленность». Геометрически это соответствует крайне вытянутым параллелепипедам, где узлы решетки расположены очень близко друг к другу вдоль определенных направлений. В такой геометрии стандартные методы линейного выравнивания приводят к катастрофическому усилению шума. Именно поэтому предлагается переход к сферическому декодированию (Sphere Decoding), которое выполняет поиск ближайшего узла внутри гиперсферы радиуса R: $|\mathbf{y} - \mathbf{H}\mathbf{x}|^2 \le R^2$. Т.о. представление задачи эквалайзинга как задачи CVP позволяет применить аппарат геометрии чисел, в частности — алгоритмы приведения базиса (LLL) и перечисления узлов (K-best), что обеспечивает достижение ML-качества детектирования при существенно меньших вычислительных затратах по сравнению с полным перебором состояний в алгоритме Витерби.

Итеративная оценка параметров канала с помощью EM-алгоритма

В условиях ДКМВ-связи импульсная характеристика канала $\mathbf{h}$ является нестационарной величиной, подверженной доплеровскому рассеянию и селективным замираниям. Для корректного функционирования сферического декодера требуется точное знание матрицы канала $\mathbf{H}$. В данной работе предлагается совместная оценка канала и детектирование данных при помощи итерационного EM-алгоритм (Expectation-Maximization). Его задача заключается в нахождении оценки максимального правдоподобия (ML) для вектора параметров канала $\mathbf{h}$, при условии, что передаваемая последовательность символов $\mathbf{x}$ является скрытой (неизвестной) переменной. Алгоритм реализует итерационный процесс, состоящий из двух чередующихся шагов: 1. E-шаг (Expectation — Ожидание). На текущей итерации j фиксируется имеющаяся оценка канала $\mathbf{h}(j)$ и вычисляется апостериорная вероятность передаваемых символов. В отличие от жесткого детектирования, результатом E-шага является вычисление мягких статистик — математического ожидания символов $\mathbf{y}$: $\hat{x}_n^{(j)} = E[x_n \mid \mathbf{y}, \mathbf{h}^{(j)}] = \sum_{s\in\mathcal{S}} s \cdot P(x_n = s \mid \mathbf{y}, \mathbf{h}^{(j)})$ Для эффективного вычисления этой суммы в пространстве решетки используется алгоритм K-best, который ограничивает перебор только наиболее вероятными узлами, попавшими в гиперсферу. Полученные на выходе LLR-метрики от нейросетевого модуля Transformer-Lite позволяют получить максимально достоверные значения $\hat{\mathbf{x}}_n$. 2. M-шаг (Maximization — Максимизация). На данном шаге обновляется оценка вектора канала $\mathbf{h}^{(j)}$ путем максимизации функции

правдоподобия на основе полученных мягких символов с помощью решения системы нормальных уравнений методом наименьших квадратов: $\mathbf{h}^{(j+1)} = \arg\min_{\mathbf{h}} \left\| \mathbf{y} - \hat{\mathbf{X}}(j)\mathbf{h} \right\|^2 = \left(\hat{\mathbf{X}}^{(j)H} \hat{\mathbf{X}}^{(j)} \right)^{-1} \hat{\mathbf{X}}^{(j)H} \mathbf{y}$, где $\hat{\mathbf{X}}^{(j)}$ — матрица данных, сформированная из оценок $x_n^{(j)}$. Использование мягких символов на этом этапе обеспечивает естественную регуляризацию: символы с низкой достоверностью (в зонах глубоких замираний) имеют малый вес и не вносят искажений в уточнение модели канала. Но для обеспечения численной устойчивости итерационного процесса на начальных этапах итеративного обновления оценки канала (М-шаг EM-алгоритма) оценка может проводиться с использованием регуляризации Тихонова: $\mathbf{h}^{(j+1)} = \left(\hat{\mathbf{X}}^{(j)H} \hat{\mathbf{X}}^{(j)} + \lambda \mathbf{I} \right)^{-1} \hat{\mathbf{X}}^{(j)H} \mathbf{y}$.

Данный подход позволяет стабилизировать процедуру обращения корреляционной матрицы данных $\mathbf{X}^H\mathbf{X}$ в условиях неполной априорной определенности символов на начальных итерациях EM-алгоритма, предотвращая расходимость оценки вектора канала $\mathbf{h}$.

Т.о. интеграция EM-алгоритма и сферического декодера превращает статическую задачу CVP в адаптивный процесс Turbo-эквалайзинга, что позволяет системе сохранять работоспособность при быстром изменении параметров многолучевости, обеспечивая стабильную работу при модуляции, нуждающейся в когерентном приёме.

Поиск ближайшего вектора в решетке большой размерности является NP-сложной задачей. Алгоритм K-best позволяет избежать полного перебора всех M^N возможных комбинаций символов. Он представляет собой субоптимальную стратегию т.н. «поиска в ширину» (Breadth-First Search), которая обеспечивает фиксированную вычислительную сложность и высокую параллельность вычислений, что критично для реализации в реальном времени. Принцип его работы следующий: алгоритм последовательно обходит дерево решений от последнего уровня $i = N$ до первого $i = 1$. На каждом уровне выполняется частичное вычисление евклидова расстояния (Partial Euclidean Distance, PED): $d_i^2(z_i, \ldots, z_N) = d_{i+1}^2 + \left| \tilde{y}i - \sum k = i^N R_{i,k} z_k \right|^2$. Ключевой особенностью K-best является то, что на каждом уровне дерева сохраняется не более K наиболее вероятных кандидатов (путей с минимальной метрикой d^2), а все остальные ветви дерева отсекаются, что позволяет радикально сократить пространство поиска. На последнем этапе алгоритм формирует список из K векторов для вычисления предварительных мягких решений (LLR), которые в дальнейшем уточняются нейросетевым модулем. Поскольку количество сохраняемых ветвей фиксировано, то алгоритм K-best, в отличии от классического алгоритма Шнорра — Эйхнера с поиском в глубину (DFS), имеет предсказуемое время работы, которое не зависит от уровня шума в канале, что упрощает проектирование как ПЛИС, так и DSP систем.

Ортогонализация базиса решетки: алгоритм Ленстры — Ленстры — Ловаса (LLL)

Как было указано ранее, эффективность поиска в решетке напрямую зависит от «качества» её базиса. Поскольку из-за корреляции столбцов матрица $\mathbf{H}$ часто является плохо обусловленной, то гиперсфера поиска превращается в сильно вытянутый эллипсоид, что вынуждает увеличивать параметр K и снижает точность детектирования. Для решения этой проблемы в структуру эквалайзера введена процедура LLL-приведения (Lattice Basis Reduction) [2]. Целью является нахождение нового базиса $\mathbf{H}_{red}$, близкого к ортогональному, который описывает ту же самую решетку, что и исходный базис. Связь между базисами $\mathbf{H}_{red} = \mathbf{HT}$ выражается через унимодулярную матрицу $\mathbf{T}$, которая состоит исключительно из целых чисел, а её определитель удовлетворяет условию $\det(\mathbf{T}) = \pm 1$, гарантирующему, что при трансформации пространство решетки и расстояние между её узлами остаются неизменными. Но применение LLL-приведения меняет саму структуру задачи детектирования: вместо поиска исходного вектора символов $\mathbf{x}$ эквалайзер ищет вектор $\mathbf{z}$ в трансформированном пространстве, поскольку выполняется трансформация принятого сигнала с использованием матрицы $\mathbf{Q}$ от приведенного базиса: $\tilde{\mathbf{y}} = \mathbf{Q}^H \mathbf{y} = \mathbf{R}_{red} \mathbf{z} + \mathbf{n}'$. Поиск вектора $\mathbf{z}$ осуществляется алгоритмом K-best в «выпрямленной» решетке $\mathbf{R}_{red}$. Благодаря почти ортогональной структуре матрицы, первый же найденный путь в

дереве с высокой вероятностью является оптимальным. После нахождения оценки $\hat{z}$ выполняется обратный переход в исходное пространство исходных символов: $\hat{\mathbf{x}} = \mathbf{T}\hat{\mathbf{z}}$.

Реализация итерационного цикла «Детектор — Оценка канала»

Для достижения максимальной помехоустойчивости в ДКМВ-канале недостаточно разового применения LLL-приведения, поэтому предлагается архитектура итеративного совместного детектирования и оценивания, где блоки LLL, K-best и EM-алгоритм образуют замкнутую петлю обратной связи. Цикл обработки одного кадра данных состоит из последовательного выполнения следующих этапов: Начальный этап: 1) Используя ограниченный набор пилот-сигналов (или чисто слепую оценку, например, методом кластеризации K-средних), вычисляется первичная оценка вектора канала $\mathbf{h}^{(0)}$. На данном этапе неопределенность велика, поэтому матрица решетки $\hat{\mathbf{H}}^{(0)}$ является лишь грубым приближением истинной. 2) Выполняется QR-разложение и LLL-ортогонализация текущей оценки матрицы $\mathbf{H}^{(j)}$. Особенность реализации в том, что при малых изменениях задержек лучей в канале (особенность ДКМВ), но даже при изменениях фаз лучей пересчет унимодулярной матрицы $\mathbf{T}$ между итерациями может быть пропущен, что существенно экономит вычислительный ресурс ПЛИС/DSP. 3) Мягкое детектирование (E-шаг): Алгоритм K-best в ортогональном базисе формирует список кандидатов. Нейросетевой модуль Transformer-Lite корректирует LLR-метрики, учитывая не только геометрию решетки, но и выявленные станционные помехи. 4) Уточнение модели (M-шаг): На основе «мягких» символов $\hat{\mathbf{x}}$, полученных из LLR, выполняется пересчет вектора $\mathbf{h}^{(j+1)}$. Поскольку в расчет принимаются все информационные символы кадра, точность оценки канала растет с каждой итерацией.

Интеграция нейросетевого модуля Transformer-Lite. Обоснование применения Deep Learning для коррекции LLR

Классические аналитические методы расчета мягких решений (LLR) в сферических декодерах базируются на допущении об аддитивном белом гауссовском шуме. Однако в реальных условиях ДКМВ-связи доминируют станционные помехи, имеющие выраженную временную и частотную структуру. Применение глубокого обучения позволяет заменить жесткие аналитические формулы адаптивной моделью, которая способна: 1) Идентифицировать сигнатуры помех, т.е. распознавать «отпечатки» нелинейных и сосредоточенных помех в векторе остаточной ошибки; 2) корректировать геометрию решетки, т.е. компенсировать искажения границ созвездия QAM/PSK, возникающие после LLL-трансформации.

Для данной задачи выбрана архитектура Transformer-Lite, представляющая собой облегченную версию трансформерной нейросети, оптимизированную для задач цифровой обработки сигналов. Ключевым преимуществом данной архитектуры является механизм Self-Attention, который позволяет сети динамически оценивать степень доверия к каждому символу внутри пакета. Математически это описывается через: вычисление матриц запросов $\mathbf{Q}$, ключей $\mathbf{K}$ и значений $\mathbf{V}$: $\text{Attention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \text{softmax}\left(QK^H / \sqrt{d_k}\right)\mathbf{V}$. В контексте эквалайзинга механизм внимания работает как адаптивный коррелятор: если на определенном интервале времени возникает станционная помеха, сеть «замечает» корреляцию ошибок между соседними символами и автоматически снижает величину LLR для данного блока, предотвращая отправку на вход декодера недостоверных данных. Для обеспечения высокой обобщающей способности сети применена гибридная стратегия обучения: для формирования обучающей выборки использована модель канала Ваттерсона для генерации искажений сигнала каналом, а в качестве аддитивных помех использовались как синтетические, так и реальные записи ДКМВ-эфира, содержащие типичные индустриальные и станционные помехи.

Для обучения нейросети в условиях ДКМВ канала со станционными помехами стандартная функция BCEWithLogitsLoss может оказаться недостаточно эффективной. Станционные помехи создают «аномальные» ошибки (outliers), которые классическая энтропия пытается минимизировать слишком агрессивно, что приводит к переобучению или нестабильности LLR. Чтобы нейросеть лучше учитывала «тяжелые хвосты» распределения помех, целесообразно использовать модифицированную функцию потерь Focal Loss [3], позволяющую сети концентрироваться на символах, пораженных помехой, предотвращая эффект усреднения,

характерный для стандартных функций потерь, что делает эквалайзер устойчивым к редким, но мощным выбросам помех: $Loss = -\alpha(1-p_t)^{\gamma}\log(p_t)$. Т.е. снизить влияние легко классифицируемых символов и сфокусировать градиент обучения на символах, искаженных сосредоточенными станционными помехами, позволив механизму Self-Attention нейросети Transformer-Lite эффективнее выделять закономерности в структуре помех с «тяжелыми хвостами» распределения. Данный подход позволяет динамически масштабировать веса градиентов, заставляя модель фокусироваться на символах, искаженных станционными помехами. В сочетании с техникой Label Smoothing это предотвращает переобучение детектора и обеспечивает получение калиброванных мягких решений (LLR), пригодных для эффективного использования в итеративной процедуре Turbo Equalization.

Сетевая архитектура и стратегия обучения сети

В известных работах, посвящённых применению DL для задач ЦОС используются, как правило, свёрточные нейронные сети (CNN), поскольку они предназначены для обработки изображений, а спектрограмму сигнала в координатах частота - время можно считать таковым. Глобальным преимуществом CNN можно считать инвариантность по отношению к частотному сдвигу сигналов и помех. Для задачи формирования LLR известны решения на основе сочетания 1D-CNN + MLP. Почему предлагается Transformer-Lite и чем это лучше сочетания 1D-CNN + MLP? 1) CNN видит только небольшое количество соседних символов, определяемое размером ядра свёртки, которое исходя из допустимой вычислительной сложности может быть только небольшим (как правило, равно 3). Transformer видит весь пакет и если станционная помеха имеет периодичность, он её обязательно обнаружит. 2) Механизм внимания работает как динамический фильтр - он «включается» только тогда, когда в данных появляется аномалия.

Таким образом, в отличии от сверточных сетей (CNN), архитектура Transformer-Lite, реализующая механизм селективного внимания, позволяет эквалайзеру учитывать долгосрочные зависимости в структуре станционных помех ДКМВ-диапазона, что обеспечивает коррекцию мягких решений (LLR) в условиях, когда классические методы детектирования (K-best) демонстрируют деградацию из-за неадекватности гауссовской модели шума.

Стратегия обучения состоит в комбинации следующих методов: 1) Label Smoothing: вместо обучения на жестких [0, 1] обучать лучше на [0.05, 0.95]. Это не даст сети стать «самоуверенной» и генерировать бесконечные LLR, что критично для последующего итеративного процесса EM; 2) SNR-Augmented Loss: можно подавать значение текущего SNR в функцию потерь, т.е. при низком SNR мы снижаем штраф, чтобы сеть не пыталась «выучить» чистый шум.

Моделирование и анализ результатов

Использовалась модель ДКМВ канала согласно ITU-R [4] и виды модуляции 16-64QAM, BPSK, QPSK, 8PSK. Предлагаемый алгоритм (LLL + K-best + EM + DL) демонстрирует преимущество в 7–9 дБ по сравнению с линейным методом выравнивания. По сравнению с классическим турбо эквалайзером предлагаемый вариант демонстрирует: 1) расширение области сходимости в сторону меньших значений сигнал/шум при наличии только шума (Рисунок 1, а); 2) существенный энергетический выигрыш при наличии станционных помех (Рисунок 1, б). Благодаря LLL-ортогонализации наклон кривой BER более крутой, что свидетельствует о более полном использовании эффекта разнесения лучей. Нейросетевой модуль Transformer-Lite обеспечивает дополнительный выигрыш в 1.5–2 дБ при действии станционных помех, успешно корректируя LLR-метрики. Установлено, что при уровне SNR выше 8 дБ алгоритм достигает стабильной оценки канала всего за три итерации. Это подтверждает высокую скорость сходимости предложенной адаптивной петли обратной связи EM алгоритма.

Несмотря на использование сложных математических процедур, алгоритм пригоден для реализации в реальном времени, поскольку: 1) фиксированная сложность K-best делает время демодуляции предсказуемым; 2) использование «легкой» архитектуры Transformer-Lite (всего 2–4 головы внимания) позволяет выполнять инференс нейросети DSP или FPGA за время, не превышающее длительности нескольких символов обрабатываемого пакета; 3) LLL-приведение решётки может выполняться не для каждого символа, а поблочно, что снижает общие затраты вычислительных мощностей на 40–60%.

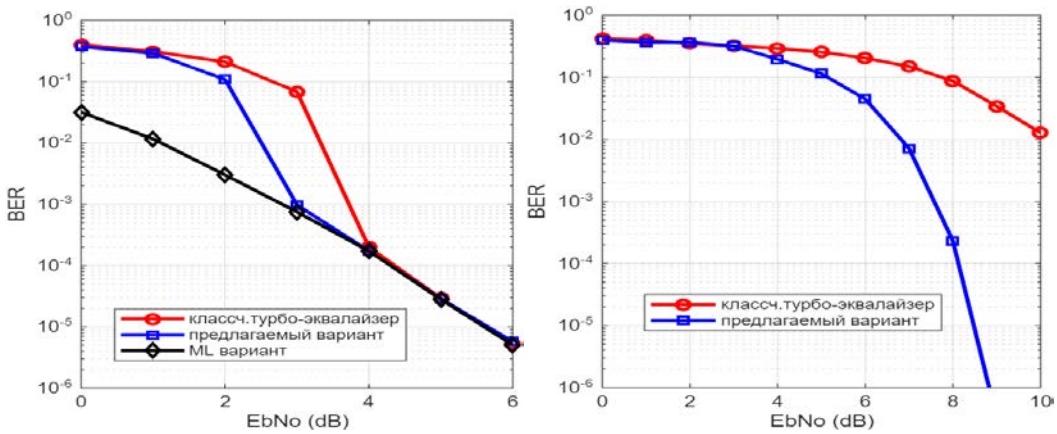


Рисунок 1. Кривые помехоустойчивости для режима 16QAM в «удовлетворительном» канале [4]. а) Приём на фоне АГБШ; б) приём на фоне АГБШ и помехи

Заключение. Основные выводы

Таким образом, был разработан итеративный эквалайзер для ДКМВ-систем связи, построенный с использованием полуслепой оценки канала на основе EM алгоритма, демодуляции символов при помощи совмещения алгоритма приведения решетки LLL типа с алгоритмом K-best нейросети, позволяющей выполнять формирование LLR бит в негауссовских шумах (включая воздействие станционных помех), характерных для ДКМВ диапазона.

Представление задачи детектирования как поиска ближайшего вектора в решетке позволило применить аппарат геометрии чисел для борьбы с многолучевостью. Использование LLL-алгоритма совместно с K-best позволило достичь помехоустойчивости, близкой к пределу ML-детектора, при сохранении полиномиальной сложности. Нейросеть обеспечила возможность адаптации эквалайзера к негауссовским станционным помехам через механизм самовнимания. Моделирование в среде MATLAB подтвердило превосходство разработанного метода над классическими линейными эквалайзерами, а также классическими турбо-эквалайзерами при наличии станционных помех, характерных для работы в условиях реального эфира.

Литература

1. Zhang R. et al. Deep learning applications in ionospheric modeling: Progress, challenges, and opportunities //Remote Sensing. – 2025. – Т. 17. – №. 1. – С. 124.
2. Wubben, D., et al. Lattice-reduction-aided detectors for MIMO communication systems. The Second International Workshop on Signal Processing for Wireless Communications, 2004.
3. Terven J. et al. A comprehensive survey of loss functions and metrics in deep learning //Artificial Intelligence Review. – 2025. – Т. 58. – №. 7. – С. 195.
4. Recommendation ITU-R F.1487. Testing of HF modems with bandwidths of up to about 12 kHz using simulated channels.

NEURAL NETWORK TURBO EQUALIZER USING DEEP LEARNING BASED ON SELF-ATTENTION MECHANISM

Malyutin A.A.[1], M.D. Portnichenko M.D.[1], A.N. Zyuzin A.N.[1], Khoroshailova M.V.[2]

[1]JSC Scientific and Innovation Enterprise "PROTEK", Voronezh, Russia
[2] Voronezh State Technical University, Voronezh, Russia

We propose a turbo equalizer with soft-decision generation based on the Transformer-Lite neural network, integrated into an iterative spherical decoding loop with the Schnorr-Eichner enumeration strategy in the K-best configuration. The neural network uses a self-attention mechanism to account for the presence of station interference in real broadcasts. Training is performed on a hybrid dataset, including both synthetic data and real broadcast recordings. The algorithm has a polynomial and strictly fixed computational complexity.

МОДЕЛЬ ШИРОКОПОЛОСНОГО MIMO/SIMO ДКМВ КАНАЛА ДЛЯ РЕАЛИЗАЦИИ МНОГОПАРАМЕТРИЧЕСКОЙ АДАПТАЦИИ, ИСПОЛЬЗУЮЩЕЙ ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ

Зюзин А.Н.

АО НВП «ПРОТЕК», г. Воронеж, Россия

В работе предложена модель широкополосного многоантенного (SIMO/MIMO) декаметрового канала, объединяющая известные модели Воглера-Хофмейера и Вексельбергера, а также учитывающая азимутальные отклонения лучей от дуги большого круга. Самой важной и отличительной особенностью модели является её дифференцируемость, введённая для возможности использования в процессе обучения сквозных нейросетей, включая сети для обучения с подкреплением с целью реализации алгоритма многопараметрической адаптации.

I. Введение

В последнее время в системах ДКМВ связи стали внедрятся широкополосные виды сигналов (ШПС с прямым расширением спектра, ППРЧ) в сочетании с современными методами многоантенного приёма (MIMO с каналом обратной связи, SIMO с адаптивным формированием диаграммы направленности (ДН) на передаче и/или на приёме с управлением по обратному каналу от Tx на Rx). Для эффективного функционирования системы подобного рода нуждаются в оптимальном выборе параметров, обеспечивающих достижение определённого показателя качества связи на основе критерия, которым может быть: 1) минимизация величины битовой ошибки (BER) при заданной скорости передачи, 2) максимальная скорость при величине BER не выше заданной или любой аналогичный критерий, вид которого определяется сценарием использования, оконечной аппаратурой и т.д. Поскольку показатель качества сложным и плохо формализуемым образом зависит от состояния канала связи, то выбор оптимальных параметров является непростой задачей. В [1] был предложен и в ряде отечественных и зарубежных систем связи реализован способ выбора оптимальных параметров на основе использования встроенного непосредственно в приёмник имитатора канала связи, реализующего модель канала, параметры которого определяются в ходе сеанса на приёмной стороне и через который пропускаются все варианты сигнально-кодовых конструкций, позволяя выбрать наилучшую для данных (оцененных приёмником) условий. Оптимальные параметры сигнально-кодовых конструкций передаются с приёмной стороны на передающую по каналу обратной связи. Поскольку ДКМВ канал характеризуется невысокой скоростью передачи информации, данная схема адаптации работоспособна в реальном времени, позволяя выбирать наилучшие параметры сигнала (тип и избыточность помехоустойчивого кода, вид модуляции и др.) при оцененных параметрах канала связи (числе лучей, характеристиках замираний, доплеровских характеристиках канала, параметрах помех и т.д.). Однако сложность метода многопараметрической адаптации, основанной на переборе всех вариантов, экспоненциально растёт с увеличением числа регулируемых параметров и с добавлением новых степеней свободы, связанных с введением пространственной обработки и расширением полосы сигнала, в итоге становится неприемлемой. В связи с этим появляется необходимость разработки какого-либо целенаправленного алгоритма поиска наилучшего сочетания параметров. В настоящее время наиболее перспективным вариантом решения данной задачи является использование методов глубокого обучения с подкреплением (Reinforcement Learning), в ходе которого частично предобученная нейросеть дообучается в процессе работы выбору оптимальных параметров системы, используя вышеописанный подход на основе модели канала, оцениваемой на приёмной стороне. Но, поскольку в ходе адаптации изменяются параметры как на стороне Rx, так и на стороне Tx, то такая модель канала должна быть дифференцируемой, т.е. обладать возможностью обратного распространения градиентов, необходимого для минимизации функции потерь как в ходе своего первоначального обучения, так и в процессе эксплуатации.

Таким образом для обучения сквозных (End-to-End) нейросетевых моделей связи в ДКМВ диапазоне требуется математическая модель канала, которая учитывает его:

1) широкополосность (в первую очередь внутрилучевое рассеяние и частотную дисперсию); 2) пространственную структуру: корреляционные свойства как антенных решеток, так и самого канала; 3) дифференцируемость: возможность вычисления градиентов по физическим параметрам канала.

Несмотря на то, что для фреймворка машинного обучения TensorFlow корпорацией NVIDIA была разработана библиотека дифференцируемых моделей каналов связи SIONNA [2,3], представленные в ней модели (изначально предназначенные только для 5G/6G) не могут быть использованы для широкополосных каналов с частотной дисперсией даже при подстановке в них численных параметров каналов, характерных для ДКМВ, а также они не обеспечивают работу в реальном времени, поскольку предназначены для использования только на этапе обучения сети.

Поэтому целесообразно разработать собственную дифференцируемую модель, учитывающую все желаемые особенности канала и работоспособную в реальном времени для реализации возможности обучения с подкреплением.

II. Выбор многоантенной (MIMO) модели ДКМВ канала

Как известно, наиболее полным способом описания MIMO канала, в том числе и коррелированного, является описание, основанное на векторизации его корреляционной матрицы. Однако сложность такого описания делает его неприемлемым. (Аналогичная проблема имеет место и при традиционном моделировании любого пространственно-коррелированного MIMO канала.) Как правило, в подавляющем большинстве работ для упрощения математического описания коррелированного MIMO канала используется модель Кронекера:

$$\mathbf{H} = \mathbf{R}_{Rx}^{1/2}\mathbf{H}\left(\mathbf{R}_{Tx}^{1/2}\right)^{H}, \tag{1}$$

где $\mathbf{H}_W$ - матрица независимых одинаково распределённых гауссовских случайных величин, а $(\cdot)^{1/2}$ - матричный квадратный корень, вычисляемый через разложение Холецкого. Полная корреляционная матрица представляет собой произведение Кронекера корреляционных матриц на сторонах Tx и Rx: $\mathbf{R}_H = E\left\{vec(\mathbf{H})vec\left(\mathbf{H}\right)^H\right\} = \mathbf{R}_{Tx}^T \otimes \mathbf{R}_{Rx}$. Модель проста в применении и интерпретации и позволяет легко рассчитать свои параметры по экспериментальным данным как: $\boldsymbol{R}_{\mathrm{Rx}} = E\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathrm{H}}\right\}$, $\boldsymbol{R}_{\mathrm{Tx}} = E\left\{\hat{\mathbf{H}}^{\mathrm{T}}\hat{\mathbf{H}}^{*}\right\}$. Однако, модель Кронекера для MIMO канала предполагает отсутствие корреляции между углами прихода и вылета сигналов на приёмной и передающей сторонах, что вполне может быть применимо для каналов сотовой и/или беспроводный связи, но никак не подходит для описания ДКМВ канала [4].

Более адекватным способом описания MIMO ДКМВ канала является модель Вексельбергера [4], которая описывается как [5]:

$$\mathbf{H} = \mathbf{U}_{Rx}\left(\tilde{\boldsymbol{\Omega}} \odot \mathbf{G}\right)\mathbf{U}_{Tx}^{T}. \tag{2}$$

Расчёт параметров модели по экспериментальным данным производится: 1) вычислением выборочных корреляционных функций $\hat{\mathbf{R}}_{\mathrm{T}x} = \frac{1}{T}\sum_{x}^{T}\boldsymbol{H}^{\mathrm{T}}\left(t\right)\boldsymbol{H}^{*}\left(t\right)$ и $\hat{\mathbf{R}}_{\mathrm{R}x} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{H}\left(t\right)\boldsymbol{H}^{\mathrm{H}}\left(t\right)$, где T - объём статистики; 2) разложений их по собственным векторам и собственным числам - $\hat{\mathbf{R}}_{Rx} = \hat{\mathbf{U}}_{Rx}\boldsymbol{\Lambda}_{xx}\hat{\mathbf{U}}_{Rx}^{H}$, $\hat{\mathbf{R}}_{\mathrm{Tx}} = \hat{\mathbf{U}}_{\mathrm{Tx}}\hat{\boldsymbol{\Lambda}}_{\mathrm{Tx}}\hat{\mathbf{U}}_{\mathrm{Tx}}^{\mathrm{H}}$; 3) вычисления вспомогательной функции $\boldsymbol{K}\left(t\right) \triangleq \hat{\mathbf{U}}_{Rx}^{H}\boldsymbol{H}\left(t\right)\hat{\mathbf{U}}_{Tx}^{H}$; 4) расчёта матрицы связи - $\hat{\tilde{\boldsymbol{\Omega}}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{K}\left(t\right)\odot\boldsymbol{K}^{*}\left(t\right)$. Матрица $\mathbf{G}$ в (2) состоит из независимых одинаково распределённых гауссовских случайных величин.

Структура матрицы связи $\tilde{\boldsymbol{\Omega}}$ отражает пространственное расположение отражающих объектов на передающей и приёмной сторонах, что, в отличии от модели Кронекера (1), позволяет учитывать влияние пространственных факторов на: 1) пропускную способность, 2) показатель разнесения, 3) число параллельных виртуальных каналов в MIMO канале, которые могут быть мультиплексированы и 4) достижимый выигрыш от диаграммообразования.

Пространственная структура многоглучевого канала для частоты f, дополнительно учитывающая возможность отклонения радиоазимута ϕ_p от дуги большого круга трассы из-за возможного наклона отражающих областей ионосферы, описывается модифицированной моделью Вексельбергера [5]:

$$\boldsymbol{H}(f,t)=\sum_{p=1}^{P}L_p(f)\cdot\left[\boldsymbol{a}_{Rx}\left(\phi_p^{(Rx)}\right)\cdot\left(\boldsymbol{\Omega}_p\odot\mathbf{G}\right)\cdot\boldsymbol{a}_{Tx}^{H}\left(\phi_p^{(Tx)}\right)\right], \quad (3)$$

где $\boldsymbol{a}_r\left(\phi_p^{(Rx)}\right)$ и $\boldsymbol{a}_t\left(\phi_p^{(Tx)}\right)$ — векторы пространственного отклика (steering vectors) антенных решеток на передающей и приёмной сторонах, а $\boldsymbol{\Omega}_p$ — матрица энергетического взаимодействия лучевых компонент.

III. Широкополосная модель ДКМВ канала

Проявление эффектов широкополосности в ДКМВ канале состоит в наличии: 1) внутрилучевого рассеяния, приводящего к размытию отклика на луч во временной области и 2) частотной дисперсии, приводящей к непостоянству группового времени задержки, вызывающем нелинейную зависимость фазы от частоты. Для учёта данных эффектов наиболее часто используется модель Воглера-Хофмейера [4,6-8], согласно которой для описания широкополосного распространения в ДКМВ диапазоне, импульсная характеристика представляется в частотной области как суперпозиция P дискретных лучей, а передаточная функция p-го луча с учетом внутрилучевого размытия и частотной дисперсии, имеет вид:

$$L_p(f)=A_p(f)\cdot\exp\left(-j\left[2\pi\tau_p(f)f+\pi\dot{\tau}_p(f)f^2+\varphi_p(t)\right]\right), \quad (4)$$

где: $\dot{\tau}_p=\partial\tau(f)/\partial f$ — коэффициент дисперсии (параметр Воглера), описывающий нелинейный фазовый набег; $\tau_p(f)$ — групповая задержка центра луча; $\varphi_p(t)$ — случайная фаза.

IV. Объединение широкополосной и многоантенной (MIMO/SIMO) моделей канала

Пространственная структура широкополосного многолучевого MIMO канала для частоты f описывается с помощью объединения модифицированной модели MIMO канала Вексельбергера (3), учитывающей возможность отклонения радиоазимута ϕ_p от дуги большого круга трассы и широкополосной модели Воглера-Хофмейера (4), учитывающей внутрилучевое рассеяние и дисперсию:

$$\boldsymbol{H}(f,t)=\sum_{p=1}^{P}L_p(f)\cdot\left[\boldsymbol{a}_{Rx}\left(\phi_p^{(Rx)}\right)\cdot\left(\boldsymbol{\Omega}_p\odot\tilde{\boldsymbol{H}}_{w,p}(t)\right)\cdot\boldsymbol{a}_{Tx}^{H}\left(\phi_p^{(Tx)}\right)\right],$$

где $\tilde{\boldsymbol{H}}(f,t)$ — стохастическая матрица отклика, элементы которой объединяют замирания и широкополосные искажения: $\tilde{\boldsymbol{H}}_{i,j}(f,t)=\sum_{p=1}^{P}\alpha_{i,j,p}\cdot L_p(f)\cdot\exp\left(j2\pi\nu_p t\right)$, а $\alpha_{i,j,p}$ — комплексный коэффициент замираний луча (обычно распределение Релея), а ν_p — доплеровский сдвиг.

V. Адаптация модели для выполнения условия дифференцируемости

Адаптация путём управления множеством параметров системы связи базируется на парадигме обучения с подкреплением, реализованной по архитектуре «Актер-Критик» (Actor-Critic). Данный подход сочетает градиентную оптимизацию стратегии выбора режимов (Актер) с нейросетевой аппроксимацией функции ценности (Критик). Под аппроксимацией функции ценности понимается способность модели предсказывать ожидаемый кумулятивный выигрыш (награду) для текущей совокупности параметров ионосферы ($\tau,\dot{\tau},\phi$). Это позволяет системе

отличать кратковременные флуктуации сигнала от устойчивых физических трендов и принимать решения, максимизирующие пропускную способность на длительном интервале времени.

VI. Дифференцируемость по задержке

Производная выходного сигнала по параметру задержки τ_p вычисляется с использованием приёма sinc-интерполяции дробной величины задержки при которой временной сдвиг τ_p дискретной последовательности $x[n]$ преобразуется в дифференцируемую операцию свертки с ядром: $x\left(t-\tau_p\right) \approx \sum_{k=-K}^{K} x\left[n-k\right] \cdot \mathrm{sinc}\left(k-\tau_p\right)$. Это позволяет вычислять производную $\partial y / \partial \tau_p$: $\frac{\partial y(t)}{\partial \tau_p} = \int x\left(t-\tau\right) \cdot \frac{\partial \mathrm{sinc}\left(\tau-\tau_p\right)}{\partial \tau_p} d\tau.$

VII. Репараметризация случайного процесса замираний

В классических моделях канала случайный характер замираний создает разрыв в вычислительном графе, что делает невозможным использование градиентных методов обучения. Репараметризация замираний это математический прием, который позволяет «пробросить» градиент через случайную величину. Этот метод используется для обучения вариационных автокодировщиков и применим в обучении с подкреплением. Для моделирования случайного процесса замираний $\alpha_{i,j,p}$ представим их в виде: $\alpha = \mu + \sigma \cdot \epsilon$, $\epsilon \sim \mathcal{CN}\left(0,1\right)$, что делает статистические параметры канала $\left(\mu, \sigma\right)$ обучаемыми, поскольку источник случайности выносится за пределы обучаемого контура. Теперь модель рассматривает случайный процесс замираний как фиксированный внешний сигнал, а параметры мощности замираний становятся обычными множителями. Это превращает стохастический канал в прозрачный дифференцируемый слой, позволяя нейросети «видеть» структуру замираний и адаптироваться к ней. Т.е. случайный канал становится дифференцируемой функцией, где параметры замираний становятся обычными обучаемыми весами, что позволяет использовать стандартный градиентный спуск.

VIII. Обеспечение дифференцируемости числа лучей

Для обеспечения дифференцируемости числа лучей используется тот же подход что и в SIONNA для реализации дифференцируемой трассировки лучей (Ray Tracing), построенный на основе математического приема Gumbel-Softmax [9], который позволяет обучать нейросети делать «дискретный выбор», сохраняя при этом возможность вычислять градиенты для обучения. Он позволяет оптимизировать дискретные параметры (например, количество лучей или выбор путей распространения) с использованием градиентных методов обучения.

Распределение Гамбеля является предельным распределением максимума последовательности независимых и одинаково распределенных случайных величин (экстремальная порядковая статистика). Плотность вероятности (PDF) стандартного распределения Гамбеля ($\mu = 0, \beta = 1$): $f\left(x\right) = \exp\left\{-x - \exp\left(-x\right)\right\}$. Для генерации шума Гамбеля g_i в коде используется датчик равномерного распределения $u_i \sim \mathrm{Uniform}\left(0,1\right)$, тогда $g_i = -\ln\left(-\ln\left(u_i\right)\right)$. Поскольку операция $\arg\max$ не имеет градиента, в SIONNA применяется «мягкая» аппроксимация (softmax) с т.н. «параметром температуры» τ: $y_i = \frac{\exp\left(\left(\ln \pi_i + g_i\right)/\tau\right)}{\sum_{j=1}^{k} \exp\left(\left(\ln \pi_j + g_j\right)/\tau\right)}$, где π_i — логиты вероятности, g_i — стандартный шум Гамбеля, τ — параметр температуры, который управляет тем, насколько «жестким» будет выбор.

IX. Описание алгоритма многопараметрической адаптации путём обучения с подкреплением на базе Actor-Critic с дифференцируемым имитатором

Взаимодействие нейросетевого агента с дифференцируемым имитатором ионосферного канала описывается следующей системой уравнений:

$$a_t \sim \pi_\theta(a_t \mid s_t), \tag{5}$$

$$y(f) = \sum_{p=1}^{P_{max}} w_p \cdot L_p(f,\tau,\dot{\tau}) \cdot A(\phi) \cdot x(f,a_t) + n, \tag{6}$$

$$\mathcal{L}_{critic} = \| V_\omega(s_t) - (r_{Tx} + \gamma V_\omega(s_t)) \|^2, \tag{7}$$

$$\nabla_\theta J(\theta) \approx E\left[\nabla_\theta \log \pi_\theta(a_t \mid s_t) \cdot \left(\frac{\partial R}{\partial \mathcal{H}} \cdot \frac{\partial \mathcal{H}}{\partial a_t} \right) \right], \tag{8}$$

здесь уравнение (5) описывает логику выбора параметров сигнала «Актером»; (6)— физическое преобразование сигнала в широкополосном MIMO-канале; (7)— процесс нейросетевой аппроксимации функции ценности «Критиком»; (8)— механизм сквозной оптимизации параметров через аналитическое вычисление градиентов имитатора. «Актер» аппроксимирует стратегию управления $\pi_\theta(a_t \mid s_t)$, где θ — обучаемые веса нейросети и на основе вектора состояния канала $s_t = [\mathrm{SNR}, \tau, \dot{\tau}, \phi, P]^T$ формирует закон распределения вероятностей для выбора вектора параметров связи a_t, который включает индекс схемы модуляции и кодирования и веса диаграммообразования. Дифференцируемая среда $\mathcal{H}$ выступает в роли «цифрового двойника» канала, преобразующего действие a_t в наблюдаемый комплексный отклик согласно (6), где $L_p(f,\tau,\dot{\tau})$ — передаточная функция Воглера-Хофмейера; $\boldsymbol{A}(\phi)$ — пространственная матрица отклика; w_p — дифференцируемые веса Gumbel-Softmax, определяющие эффективное число лучей P; $\boldsymbol{x}(f,a_t)$ — спектр сигнала, сформированный «Актером». Имитатор вычисляет награду r_t, характеризующую эффективность выбора режима работы: $r_t = \eta \cdot \mathrm{Rate}(a_t) - \alpha \cdot \log(\mathrm{BER}(\boldsymbol{y}, \boldsymbol{x}))$, где Rate — скорость передачи. Блок «Критик» (сеть V_ω) минимизирует ошибку предсказания награды: $\mathcal{L}_{critic} = \| V_\omega(s_t) - (r_t + \gamma V_\omega(s_{t+1})) \|^2$. Благодаря дифференцируемости имитатора $\mathcal{H}$ по всем физическим параметрам, обновление весов «Актера» происходит с использованием аналитического градиента: $\nabla_\theta J(\theta) \approx \mathbb{E}\left[\nabla_\theta \log \pi_\theta(a_t \mid s_t) \cdot \left(\frac{\partial R}{\partial \mathcal{H}} \cdot \frac{\partial \mathcal{H}}{\partial a_t} \right) \right]$. Слагаемое в скобках представляет собой чувствительность функции награды к параметрам сигнала.

X. Заключение и выводы

1. Предложенный в работе синтез аналитической модели Воглера–Хофмейера и пространственной модели Вексельбергера позволяет корректно описывать частотно-селективные замирания и азимутальные отклонения радиолучей в ионосфере с учётом внутрилучевого рассеяния и частотной дисперсии.
2. В отличие от существующих библиотек имитационного моделирования [7,8,9], ориентированных на стандарты 5G/6G и тропосферное распространение, разработанный программный модуль на базе библиотеки PyTorch реализует канал как непрерывный дифференцируемый оператор. Это устраняет проблему «разрыва градиента» и открывает возможность применения градиентных методов End-to-End. Реализация позволяет исследовать

чувствительность системы именно к параметрам Воглера-Хофмейера в чистом виде, без влияния избыточных протокольных ограничений стандартов 5G/6G, заложенных в архитектуру SIONNA.
3. Использование механизмов интерполяции для задержек лучей и репараметризации стохастических замираний делает модель пригодной для использования алгоритмов обратного распространения ошибки.
4. Предложенный подход позволяет сократить время сходимости алгоритмов адаптации по сравнению с классическими методами случайного поиска поскольку предложенная реализация обладает меньшей вычислительной избыточностью по сравнению с универсальными фреймворками. Это позволяет интегрировать её непосредственно в контур управления связью в реальном времени, реализующий механизм многопараметрической адаптации, построенный с использованием методов глубокого обучения с подкреплением, обеспечивающий выбор оптимальных схем модуляции и кодирования в условиях быстро меняющейся геофизической и помеховой обстановки.
5. Предложена модель канала с динамически изменяемым числом лучей, где переход между состояниями многолучевости реализован через дифференцируемую аппроксимацию распределения Гамбеля, позволяет нейросети обучаться с учётом структурной сложности модели канала. Данный подход позволяет трансформировать дискретную задачу выбора топологии канала в непрерывную оптимизационную задачу. Это обеспечивает возможность градиентного обучения системы в условиях динамического появления и исчезновения лучей, например, при переходе через критическую частоту слоя f0F2.

Литература

1. Малютин А.А. и др. Патент № 101298 U1 Российская Федерация, МПК H04B 7/00. Имитатор КВ канала с ППРЧ : № 2010123573/09 : заявл. 09.06.2010 : опубл. 10.01.2011
2. J. Hoydis et al., "Sionna RT: Differentiable Ray Tracing for Radio Propagation Modeling," 2023 IEEE Globecom Workshops (GC Wkshps), Kuala Lumpur, Malaysia, 2023, pp. 317-321, doi: 10.1109/GCWkshps58843.2023.10465179.
3. Aoudia F. A. et al. Sionna rt: Technical report //arXiv preprint arXiv:2504.21719. – 2025.
4. Малютин А.А. и др. Об эффективности MIMO технологий в системах ДКМВ радиосвязи //REDS: Телекоммуникационные устройства и системы. – 2017. – Т. 7. – №. 4. – С. 426-430.
5. Weichselberger W. et al. A stochastic MIMO channel model with joint correlation of both link ends //IEEE Transactions on wireless Communications. – 2006. – Т. 5. – №. 1. – С. 90-100.
6. Vogler L. E., Hoffmeyer J. A. A model for wideband HF propagation channels //Radio Science. – 1993. – Т. 28. – №. 06. – С. 1131-1142.
7. Yang R., Bai Y., Hu Z. Designand Implementation of a Novel Wideband HF Communication System Based on NC-OFDM and Probabilistic Shaping //Sensors. – 2025. – Т. 25. – №. 17. – С. 5596.
8. Lv M. et al. High-Frequency Channel Modeling Based on the Multi-Source Ionospheric Assimilation Model //Remote Sensing. – 2022. – Т. 14. – №. 17. – С. 4133.
9. R. Pegurri, F. Linsalata, E. Moro, J. Hoydis and U. Spagnolini, "Toward Digital Network Twins: Integrating Sionna RT in ns-3 for 6G Multi-Rat Networks Simulations," IEEE INFOCOM 2025 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), London, United Kingdom, 2025, pp. 1-6,

A MODEL OF A WIDEBAND MIMO/SIM HF CHANNEL FOR MULTIPARAMETER ADAPTATION USING REINFORCEMENT LEARNING

Zyuzin A.N.

JSC Scientific and Innovation Enterprise "PROTEK", Voronezh, Russia

This paper proposes a model for a wideband SIMO/MIMO HF channel that combines the well-known Vogler-Hofmeyer and Weichselberger models and takes into account the azimuthal deviations of the beams from the great-circle arc. The most important and distinctive feature of the model is its differentiability, introduced to enable its use in the training process of end-to-end neural networks, including reinforcement learning networks for implementing a multiparameter adaptation algorithm.

ВОССТАНОВЛЕНИЕ И УДЕРЖАНИЕ ОПОРНОГО СИГНАЛА ДЛЯ СИМВОЛЬНОЙ СИНХРОНИЗАЦИИ УЗКОПОЛОСНЫХ СИСТЕМ ПЕРЕДАЧИ ДАННЫХ

доц. Овинников А. А.[1], студ. Спицын Д. О.[1]

[1] ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина»

В работе рассматривается задача восстановления и удержания опорного сигнала для символьной синхронизации в узкополосных системах передачи данных при низком отношении сигнал-шум. Предложен разомкнутый алгоритм, в котором вместо традиционного дифференциатора используется фильтр Гильберта, импульсная характеристика которого квазисогласована с фронтом символа. Тактовая компонента выделяется специализированным узкополосным БИХ-фильтром «Шпилька» (peaking filter).

Метод сравнивался с классической схемой на основе дифференциатора путём моделирования в Simulink при отношении сигнал-шум 0 дБ и случайной задержке символов. Показано, что амплитуда на выходе фильтра «Шпилька» нарастает в 6,9 раза быстрее. Интенсивность дребезга снижена в 2,8 раза по длительности и в 1,85 раза по числу переходов через нуль. При цифровом отношении сигнал-шум 10 дБ в предлагаемой схеме дребезг полностью отсутствует.

Полученные результаты подтверждают преимущество разработанного алгоритма по скорости захвата и надёжности удержания синхронизации в условиях сильного шума

Введение

В современных цифровых системах связи задача восстановления тактового сигнала остаётся одной из ключевых при обеспечении надёжной передачи данных. Особую актуальность данная проблема приобретает в узкополосных системах с низким энергетическим потенциалом, где время установления синхронизации и помехоустойчивость алгоритмов непосредственно влияют на пропускную способность канала.

Среди существующих подходов к символьной синхронизации разомкнутые методы привлекают исследователей относительной простотой реализации и отсутствием необходимости в обратной связи. Однако классические разомкнутые схемы демонстрируют ограниченную эффективность в условиях сильного шума, что проявляется в значительных задержках при установлении синхронизации и нестабильности фазы восстановленного тактового сигнала.

В данной работе рассматривается альтернативный подход к построению разомкнутой системы символьной синхронизации, основанный на использовании фильтрации, согласованной с формой фронта принимаемого сигнала. Проведено сравнительное исследование предлагаемого метода и классической схемы методом компьютерного моделирования.

Теоретическая часть

Символьная (временная) синхронизация — это процесс выделения моментов оптимального отсчёта принимаемых символов в цифровых системах связи. Существует два основных типа алгоритмов символьной синхронизации: разомкнутые и замкнутые [1].

Разомкнутые алгоритмы выполняют оценку параметров временной структуры сигнала и формируют тактовый сигнал без использования обратной связи (рисунок 1, а). Их идея заключается в выделении из сигнала спектральной компоненты, соответствующей частоте следования символов, с точностью до постоянного временного сдвига [2].

Замкнутые алгоритмы реализуют контур слежения, в котором локальный генератор тактовых импульсов синхронизируется с поступающим сигналом посредством сравнения локального и принимаемого сигналов, оценки ошибки синхронизации и её компенсации (рисунок 1, б) [2].

Хотя замкнутые алгоритмы обеспечивают более высокую точность, их реализация существенно сложнее. В дальнейшем рассматриваются разомкнутые методы синхронизации. Подобные схемы эффективнее всего работают в том случае если на их вход подаётся меандр с фиксированным периодом.

Существует три классических варианта разомкнутой синхронизации [1].

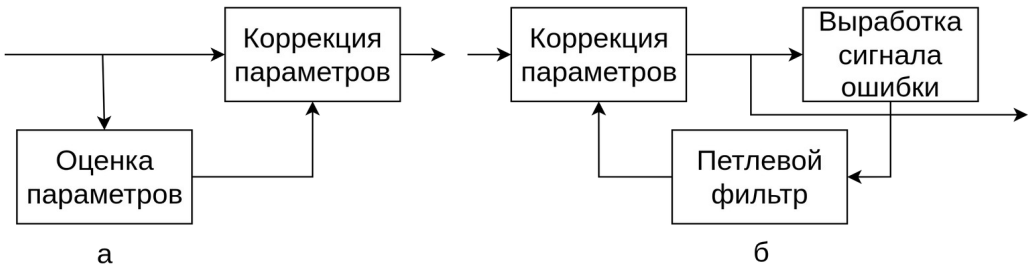


Рисунок 1 – Разомкнутая (а) и замкнутая (б) системы синхронизации



Рисунок 2 – Синхронизация на основе согласованного фильтра

Синхронизация на основе согласованного фильтра

Коэффициенты согласованного фильтра определяются формой входного сигнала. На выходе формируется последовательность импульсов, совпадающая по форме с автокорреляционной функцией входного сигнала. Для формирования однополярных импульсов с удвоенной тактовой частотой применяется нелинейность чётного порядка, в результате чего в спектре возникает компонента на удвоенной частоте следования символов. Полосовой фильтр выделяет данную компоненту, после чего блок sgn формирует периодическую последовательность прямоугольных видеоимпульсов y(t), являющуюся тактовым сигналом.

На вход схемы (рисунок 2) подаётся сигнал s(t) , представляющий собой сумму периодической последовательности биполярных прямоугольных видеоимпульсов и аддитивного белого гауссовского шума канала.

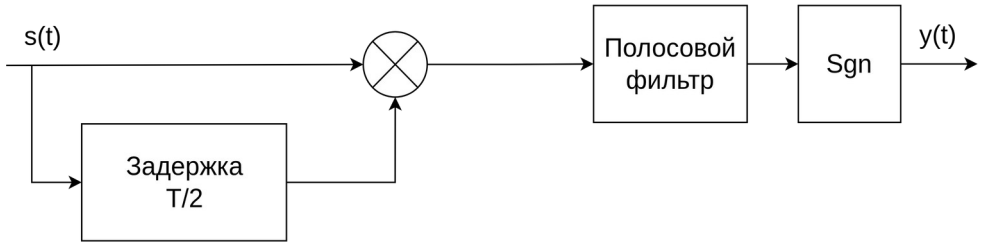


Рисунок 3 – Синхронизация на основе линии задержки

Синхронизация на основе задержки

Входной сигнал s(t) задерживается на половину длительности символа и перемножается с незадержанным. В результате первая половина символа становится положительной, а вторая — отрицательной. Задержка T/2 обеспечивает максимальную амплитуду спектральной составляющей на тактовой частоте. Далее необходимая частотная компонента, как и прежде, выделяется с помощью полосового фильтра, после чего блоком sgn формируется меандр y(t) (рисунок 3).

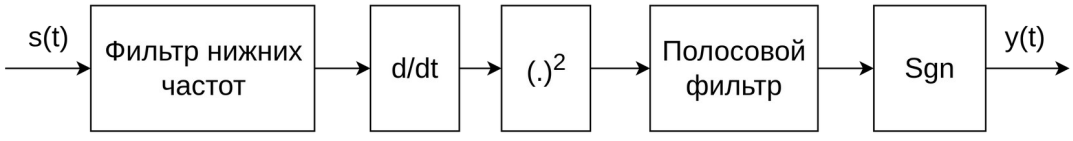


Рисунок 4 – Синхронизация на основе дифференциатора

Синхронизация на основе дифференциатора

Дифференциатор формирует резкие положительные и отрицательные выбросы на границах символов. Поскольку дифференцирование усиливает высокочастотный шум, перед ним включён фильтр нижних частот. Возведение в квадрат преобразует биполярные выбросы в однополярные импульсы, из которых полосовой фильтр выделяет спектральную компоненту на тактовой частоте, а блок sgn формирует меандр y(t) (рисунок 4).



Рисунок 5 – Предлагаемый алгоритм синхронизации

Представленные выше алгоритмы решают общую задачу — выделение моментов смены символов, однако реализуют её ра зличными способами. Предлагаемый метод отличается использованием фильтра, импульсная характеристика которого согласована с формой фронта символа. В отличие от дифференциатора, требующего предварительной фильтрации, сопровождающейся искажением формы фронта и потерей энергии сигнала, такой фильтр формирует на выходе последовательность импульсов, соответствующих границам символов, без дополнительного подавления шума. Для сигналов с прямоугольной огибающей в качестве такого фильтра предлагается применить фильтр Гильберта, чья импульсная характеристика квазисогласована с функцией Хевисайда. Далее с помощью специализированного узкополосного БИХ-фильтра «Шпилька» (peaking filter) выделяется тактовая частотная компонента, из которой блоком sgn формируется последовательность прямоугольных видеоимпульсов y(t) (рисунок 5).

Экспериментальная часть

Целью эксперимента является сравнение времени установления синхронизации и помехоустойчивости классической схемы на основе фильтра-дифференциатора и предлагаемой схемы на основе фильтра Гильберта. Ключевым критерием служит дребезг — многократное пересечение нуля выходным сигналом на начальном этапе, приводящее к ошибкам синхронизации.

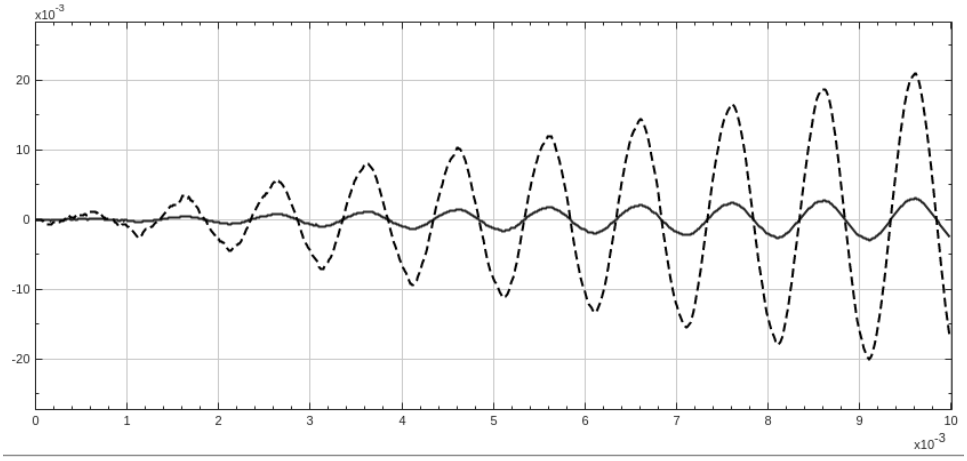


Рисунок 6 – Выход фильтра «Шпильки» двух схем
Сплошная линия - фильтр-дифференциатор
Пунктирная линия — фильтр Гильберта

Для объективного сравнения обе схемы приведены к единой структуре: различается только замена каскада ФНЧ и дифференциатора на фильтр Гильберта. Из классической схемы

исключено возведение в квадрат, поскольку оно лишь усиливает шум без улучшения выделения фронтов при данной конфигурации. Параметры фильтров: дифференциатор — вектор частот [0, 0.5, 0.55, 1], вектор амплитуд [0, 1, 0, 0], 10-й порядок; ФНЧ — граничная частота полосы пропускания 0.3, полосы заграждения 0.5, 10-й порядок; фильтр Гильберта — вектор частот [0.05, 0.95], вектор амплитуд [1, 1], 10-й порядок (в нормированных единицах к частоте дискретизации). Входной сигнал имеет частоту повторения импульсов 1 кГц, частоту дискретизации 100 кГц, центральная частота фильтра «Шпилька» (peaking filter) соответствует частоте следования импульсов. Моделирование проведено при цифровом отношении сигнал-шум 0 дБ со случайной начальной задержкой символов (jitter) с равномерным законом распределения, усреднение выполнено по 100 реализациям первых 10 периодов длительностью 10 мкс.

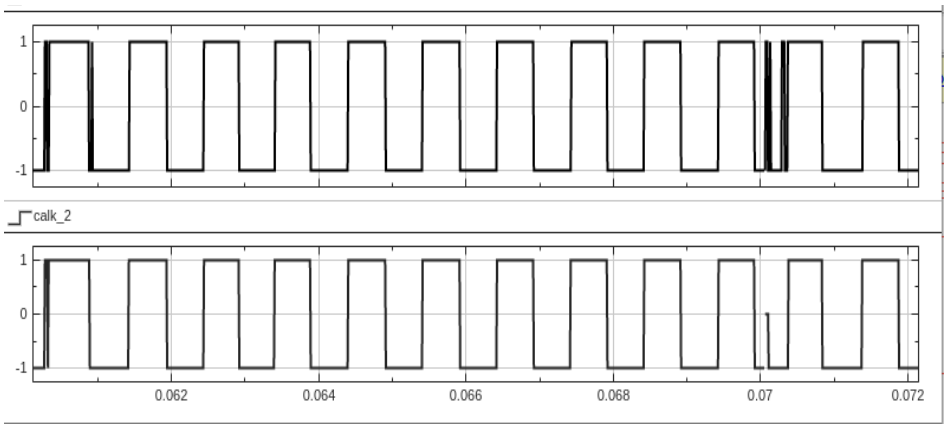


Рисунок 7 – Выходной сигнал блока sgn при E_b/N_0 0 дБ:
верхний — фильтр-дифференциатор, нижний — фильтр Гильберта

Результаты моделирования показывают, что скорость нарастания амплитуды огибающей на выходе фильтра «Шпильки» в схеме с фильтром Гильберта превышает аналогичный показатель для схемы на основе дифференциатора в 6.9 раз (рисунок 6).

На рисунке 7 представлены две реализации выходного сигнала (от 0.06с — одна реализация, от 0.07с — другая реализация) на выходе блока sgn. В схеме на основе дифференциатора (верхний график) наблюдается интенсивный дребезг в обеих реализациях, тогда как в схеме с фильтром Гильберта (нижний график) дребезг присутствует только в первой реализации, а во второй переходы границ символов чёткие.

Количественная оценка подтверждает существенное снижение интенсивности дребезга схемы на основе фильтра Гильберта относительно схемы на основе дифференциатора: средняя длительность составляет 17.95 % от длительности символа против 50.58 %, максимальная длительность — 58 % против 170 %, среднее число переходов через нуль — 3.3 против 6.1, максимальное число переходов — 5 против 13 (таблица 1). При увеличении отношения сигнал-шум до 10 дБ схема с фильтром Гильберта работает без дребезга, тогда как у дифференциатора дребезг сохраняется в каждой реализации.

Таблица 1. Значения параметров при случайной задержке и E_b/N_0 0 дБ

Название оцениваемого параметра	Значение для схемы на основе дифференциатора	Значение для схемы на основе фильтра Гильберта
Средняя длительность дребезга в процентах от длительности символа	50.58 %	17.95 %

Продолжение таблицы 1.

Максимальная длительность дребезга в процентах от длительности символа	170 %	58 %
Средняя длительность дребезга в переходах через нуль	6.1 раз	3.3 раз
Максимальное количество переходов через нуль	13 раз	5 раз
Среднее количество дребезгов за реализацию	1.4 раз	1 раз
Максимальное количество дребезгов за реализацию	2 раз	1 раз

Заключение

Предложен разомкнутый алгоритм символьной синхронизации, в котором вместо дифференциатора используется фильтр Гильберта, а тактовая компонента выделяется узкополосным БИХ-фильтром «Шпилька».

По результатам моделирования в Simulink при отношении сигнал-шум 0 дБ и случайной задержке символов амплитуда на выходе фильтра «Шпилька» в предложенной схеме нарастает в 6,9 раза быстрее, чем в схеме на основе дифференциатора.

Интенсивность дребезга (многократных пересечений нуля) в предложенной схеме ниже: средняя длительность дребезга составляет 17,95 % длительности символа против 50,58 %, среднее число переходов через нуль — 3,3 против 6,1.

При отношении сигнал-шум 10 дБ в схеме с фильтром Гильберта дребезг отсутствует во всех реализациях.

Литература

1. Скляр Б. Цифровая связь. Теоретические основы и практическое применение / Пер. с англ. — 2-е изд., испр. — М.: Вильямс, 2003. — 1104 с.
2. Сергиенко А.Б. Цифровая связь. Учебное пособие. — СПб.: СПбГЭТУ «ЛЭТИ», 2012. — 208 с.
3. Филимонов В.А., Усов Н.А., Остроумов О.А. Алгоритмы получения символьной синхронизации при использовании цифровой обработки сигналов // I-methods. 2023. № 3.

ACQUISITION AND TRACKING OF REFERENCE SIGNAL FOR SYMBOL SYNCHRONIZATION IN NARROWBAND DATA TRANSMISSION SYSTEMS

Ovchinnikov A.A.[1], Spitsyn D.O.[1]
[1] Ryazan State Radio Engineering University named after V.F. Utkin

The paper addresses the problem of reference signal acquisition and tracking for symbol synchronization in narrowband data transmission systems under low signal-to-noise ratio conditions. An open-loop algorithm is proposed in which a Hilbert filter, whose impulse response is quasi-matched to the symbol edge, replaces the conventional differentiator. The clock component is extracted using a specialized narrowband IIR peaking filter.

The method was compared with the classical differentiator-based scheme through Simulink modeling at Eb/N0 = 0 dB with random symbol delay. It is shown that the amplitude at the peaking filter output grows 6.9 times faster. Jitter intensity (multiple zero crossings) is reduced by a factor of 2.8 in duration and 1.85 in the number of zero crossings. At Eb/N0 = 10 dB, jitter is completely absent in the proposed scheme.

The obtained results confirm the advantages of the developed algorithm in acquisition speed and synchronization tracking reliability under strong noise conditions.

АСИМПТОТИЧЕСКИЕ ГРАНИЦЫ СРЕДСТВ ПЕРЕСТАНОВОЧНОГО ДЕКОДИРОВАНИЯ И ИХ АДАПТИВНАЯ ИНТЕРПРЕТАЦИЯ

доц. Ганин Д.В.

ГБОУ ВО «Нижегородский государственный инженерно-экономический университет»

Повышенные требования к скорости обработки цифровой информации в современных телекоммуникационных системах уровня 5G и выше потребовали применения в таких системах коротких избыточных кодов, чтобы уложиться в требования к задержке данных в системе связи не превышающей 1 мс. Естественно, что с уменьшением длины кодовых векторов избыточных кодов, применяемых для защиты цифровых данных от ошибок, объективно снижаются возможности таких кодов по исправлению ошибок. В указанных условиях возникает целесообразность применения средств перестановочного декодирования, которые обладают повышенной корректирующей способностью относительно известных алгоритмов повышения достоверности данных. При этом, как правило, не рассматривается ряд негативных факторов отрицательно влияющих на возможности перестановочных декодеров на достижение конечного результата. Например, появление ошибок первого и второго рода в оценке мягких решений символов кодовых векторов, или предельные возможности метода распространения доверия в алгоритмах повышения индексов мягких решений, или степень влияния методов модуляции цифровых данных на возникновение указанных ошибок. В докладе дается комплексная оценка негативных факторов, влияющих на результат перестановочного декодирования, показываются возможности адаптивной обработки данных в таких системах.

Оценка последствий ошибочной регистрации надежных символов

Замечательной особенностью блоковых систематических избыточных кодов является возможность формирования на основе одного любого из таких кодов определенного множества эквивалентных кодов (ЭК) [1, 2]. Такие коды будут иметь относительно исходного варианта одинаковые показатели длины кодового вектора n, числа информационных k разрядов, а также весовой спектр. Однако для одинаковых комбинаций информационных разрядов исходного варианта и ЭК разницей будет являться система проверочных уравнений и, следовательно, структура проверочной матрицы $H_{\text{эк}}$, которая является частью порождающей матрицы $G_{\text{эк}}$, нового кода. Алгоритм формирования такого кода достаточно прост: за счет надежных мягких решений двоичных разрядов принятого из канала связи вектора $n_{\text{кс}}$ формируется новый переставленный блок информационных разрядов $k_{\text{пер}}$, который далее за счет процедуры $k_{\text{пер}} \times G_{\text{эк}}$ формирует вектор ЭК $n_{\text{эк}}$. Сравнение упорядоченных векторов $n_{\text{эк}}$ и $n_{\text{кс}}$ позволяет определить вектор локаторов ошибок e. Поскольку для формирования ЭК требуется ровно $k_{\text{пер}}$ бит, то метод перестановочного декодирования позволяет в общем случае исправить ровно $(n - k)$ ошибок [3, 4]. Этот показатель превосходит значение метрики Хэмминга, что подчеркивает преимущество метода ЭК по реализации процедуры исправления ошибок.

Вместе с этим, в известных работах, посвященных исследованию ЭК в системах обмена данными, нет сведений о влиянии ошибок в оценке мягких решений на общий исход декодирования данных. Ошибки не в плане неправильного определения мягкого решения, а в плане сочетания высокого значения мягкого решения и соответствия этому значению собственно битовой ошибки. Если принять наилучшую оценку надежности $\lambda_{max} = 7$, то ошибкой второго рода следует считать ситуацию, когда жесткое решение принято приемником ошибочно. В этом случае будем считать, что наихудшей оценкой надежности будет оценка $\lambda_{min} = 0$ и тогда ошибкой первого рода будет безошибочное определение жесткого решения. Статистическая теория систем говорит об взаимозаменяемости данных критериев. Первоначальная интерпретация оказывается полезной и более важной с точки зрения реализации метода обработки ЭК.

Действительно, если декодер в ходе сортировки оценок λ_i на длине кодового вектора находит оценку $\lambda_{max} = 7$, он с вероятностью близкой к единице переводит этот символ в разряд $k_{\text{пер}}$.

Утверждение. Одиночный ошибочный символ с высоким показателем надежности, включенный в ходе сортировки нумераторов в группу информационных разрядов для решения задачи перестановочного декодирования, однозначно приводит к ошибочному решению по принятой комбинации при сравнении векторов $n_{\text{эк}}$ и $n_{\text{кс}}$.

Рассмотрим пример, в котором переданная кодовая комбинация основного кода (7, 4, 3) **1**010011 в ходе передачи по каналу связи в качестве надежных нумераторов получила

перестановку **1**267 (жирным выделена ошибка с высокой оценкой надежности). При отсутствии помех переставленный вектор с учетом $G_{эк}$ принял бы вид 1011100, пусть ошибка, как указывалось выше 11100е, в такой комбинации произошла на первом месте. Тогда переставленный вектор принимает вид 0011100. Тогда при поразрядном сравнении $n_{кс} =$ **1**010011 с $n_{эк} = 0011100$ локаторы ошибок указали бы на 4, 5, 6, и 7 нумератор, что далеко не так.

В количественном отношении справедливость данного утверждения проверялась с использованием аналитической модели гауссовского канала связи в условиях различных отношений сигнал-шум для формата противоположных сигналов представлены в таблице.

Таблица 1. Результаты оценки вероятности ошибок второго рода

Параметры	Показатели мягких решений							
λ_i	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**
Отношение сигнал/шум 0 дБ	3,2 10^{-3}	4,2 10^{-3}	5,5 10^{-3}	7,1 10^{-3}	9,0 10^{-3}	1,1 10^{-2}	1,4 10^{-2}	1,7 10^{-2}
Отношение сигнал/шум 0,5 дБ	2,4 10^{-3}	3,4 10^{-3}	4,5 10^{-3}	6,0 10^{-3}	7,8 10^{-3}	1,0 10^{-2}	1,3 10^{-2}	1,6 10^{-2}
Отношение сигнал/шум 1,0 дБ	1,8 10^{-3}	2,5 10^{-3}	3,5 10^{-3}	4,9 10^{-3}	6,6 10^{-3}	8,7 10^{-3}	1,5 10^{-2}	1,9 10^{-2}
Отношение сигнал/шум 3,0 дБ	2,8 10^{-4}	5,0 10^{-4}	8,6 10^{-4}	1,4 10^{-3}	2,3 10^{-3}	3,6 10^{-3}	5,4 10^{-3}	8,0 10^{-3}
Отношение сигнал/шум 5,0 дБ	9,4 10^{-6}	2,4 10^{-5}	6,0 10^{-5}	1,4 10^{-4}	2,3 10^{-3}	3,0 10^{-4}	1,3 10^{-3}	2,4 10^{-3}

При этом вероятность ошибки на бит оценивалась каноническим выражением

$$p_b = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{a1}^{a2} \exp \frac{-(z+\sqrt{2E_b})^2}{2\sigma^2} dz,$$

где σ^2 – дисперсия белого гауссовского шума, E_b – энергия сигнала на бит, z – параметр обрабатываемого сигнала, значения a1 и a2 – пределы интегрирования для зоны сигнала, определяющие формирования тех или иных значений λ_i. В работе для всех λ_i выбирались равные участки интегрирования, хотя эти участки могут быть адаптированы к уровням фиксируемых z. Интервал интегрирования $(a2 - a1) = \sqrt{E_b}/8$ при $\lambda_{max} = 7$.

Вероятность ошибочного приема комбинации $P_{км}$ с учетом применения в системе алгебраического метода исправления ошибок с использованием метрики Хэмминга в канале с независимым потоком ошибок определялась в соответствии с биномиальным законом.

$$P_{км} = \sum_{j=d_{min}}^{n} \frac{n!}{j!\,(n-j)!} (p_b)^j\,(1-p_b)^{n-j}.$$

Следует отметить, что в общей теории связи вопрос о применении того иди иного избыточного кода во многом определяется параметром энергетического выигрыша кода (ЭВК). По этой причине параметр $P_{км}$ пересчитывается в значение ошибки на бит в соответствии с выражением

$$p_{bкм} \approx \frac{1}{n} P_{км}.$$

Анализ результатов, приведенных в таблице убедительно показывает, что при отношении сигнал-шум менее 5 дБ метод ЭК становится контрпродуктивным, поскольку в число отобранных символов для создания надежной конфигурации информационных разрядов с высокой вероятностью могут попасть ошибочные символы, что немедленно приводит к ошибочному результату при формировании ЭК.

В условиях низких отношений сигнал-шум метод перестановочного декодирования без дополнительных мер по причине роста ошибок второго рода может оказаться не эффективным из-за увеличения вероятности $P_{км}$, что в итоге отрицательно скажется на значении $p_{bкм}$ и финальном ЭВК.

Оценка эффективности метода распространения доверия

Метод распространения доверия может быть использован в системе перестановочного декодирования [5, 6]. Возможности метода достаточно хорошо исследованы, его аналитическое представление имеет вид:

$$Ld_1 \boxplus Ld_2 \approx (-1)^{1+f} \times \text{sign}(Ld_1) \times \text{sign}(Ld_2) \times \min[|Ld_1|, |Ld_1|],$$

где обозначение $\boxplus$ – символизирует действия по учету знаков и абсолютных значений МРС символов Ld_1 и Ld_2; при этом значение Ld_1, по сути, представляет алгебраическое сложение некоторой априорной оценки и соответствующей ей апостериорной оценки МРС на очередном шаге итерации, т.е. $Ld_1=(d_{апр}) + (d_{апос})$ как статистически независимые информационные биты. Значение f определяется числом удаленных единиц из конкретного проверочного уравнения, определяющего проверку четности результирующего бита по $mod\ 2$. Важно, что для коррекции ошибок целесообразно иметь высокие значения мягких решений символов для проверки четности.

Метод синхронного накопления данных

При очень низких отношениях сигнал-шум в работе доказывается целесообразность многократной передачи данных с последующим применением метода мажоритарного декодирования или совместного метода синхронного накопления данных с последующим применением метода распространения доверия.

Литература

1. Питерсон У., Уэлден Э. Коды, исправляющие ошибки. М.: «Мир», 1976. 594 с.
2. Morelos-Zaragoza R. The Art of noise-tolerant coding. Methods, algorithms, and applications. Per. from the English ed. Afanasiev V.B. M.: Technosphere. 2005. 320 p.
3. Перестановочное декодирование с системой адаптированных альтернативных решений / А. А. Гладких, А. А. Овинников, Н. А. Пчелин, А. А. Брынза // Цифровая обработка сигналов. 2023. № 4. С. 73 – 78.
4. Структура и взаимосвязь когнитивных показателей в системе перестановочного декодирования / А. А. Брынза, А. А. Гладких, А. А. Ничунаев [и др.] // Автоматизация процессов управления. 2023. № 4(74). С. 126 – 133.
5. Оценка эффективности защиты данных от ошибок на базе байесовского вывода в системе итеративных преобразований / А. А. Гладких, С. М. Наместников, А. В. Новоселов [и др.] // Автоматизация процессов управления. 2022. № 4(70). С. 120 – 130.
6. Морфологический синтез в системе алгоритмов перестановочного декодирования с альтернативными решениями / С. А. Агеев, А. А. Гладких, А. А. Брынза [и др.] // T-Comm: Телекоммуникации и транспорт. 2024. Т. 18, № 3. С. 4 – 10.

ASYMPTOTIC LIMITS OF PERMUTATION DECODING TOOLS AND THEIR ADAPTIVE INTERPRETATION

docent Ganin D.V.

Nizhny Novgorod State University of Engineering and Economics

The increased demands on the processing speed of digital information in modern telecommunications systems of the 5G level and higher required the use of short redundant codes in such systems in order to meet the requirements for a data delay in a communication system of no more than 1 ms. Naturally, as the length of the redundant code vectors used to protect digital data from errors decreases, the error correction capabilities of such codes objectively decrease. Under these conditions, it becomes advisable to use permutation decoding tools, which have an increased corrective ability relative to well-known algorithms for increasing data reliability. At the same time, as a rule, a number of negative factors that negatively affect the ability of permutation decoders to achieve the final result are not considered. For example, the occurrence of errors of the first and second kind in the evaluation of soft solutions for code vector symbols, or the limiting possibilities of the method of spreading trust in algorithms for increasing the indexes of soft solutions, or the degree of influence of digital data modulation methods on the occurrence of these errors. The report provides a comprehensive assessment of the negative factors affecting the result of permutation decoding, and shows the possibilities of adaptive data processing in such systems.

МЕТОД ФОРМИРОВАНИЯ ПЕРЕСТАНОВОК И КОРРЕКЦИИ ВЫРОЖДЕННЫХ МАТРИЦ ПЕРЕСТАНОВОЧНОГО ДЕКОДЕРА

асп. Потапова С.Е., асп. Уласюк Т.Г.

Ульяновский институт гражданской авиации

В докладе рассматривается метод перестановочного декодирования коротких блоковых помехоустойчивых кодов, который обеспечивает приемлемую достоверность данных в условиях достаточно низких отношений сигнал/шум. Это происходит за счет повышения кратности, исправляемых кодом ошибок, которая, как правило, равняется числу избыточных символов. Недостатком указанного метода применительно к двоичным избыточным кодам является высокая вероятность перестановок нумераторов символов принятой кодовой комбинации, которые формируют вырожденную переставленную матрицу, без коррекции которой получить эквивалентный код становится невозможным. Для достижения положительного результата предлагается использовать метод распространения доверия, который позволяет осуществить требуемую коррекцию именно тех нумераторов предполагаемых к перестановке символов, которые наиболее рациональным образом позволяют получить эквивалентный код.

Описание проблемы

Технология перестановочного декодирования (ПД) достаточно подробно описана в работах [1–4]. Рациональным предложением для блоковых систематических кодов является предварительная оценка текущей перестановки нумераторов символов принятого кодового вектора с последующей оценкой вырожденности переставленной матрицы формируемого эквивалентного кода (ЭК). В случае невырожденной переставленной матрицы кода (тип перестановки 1) переставленный вектор направляется для дальнейшей обработки декодером. В случае вырожденной переставленной матрицы (тип перестановки 2) текущая перестановка направляется на коррекцию последовательности нумераторов символов вплоть до получения перестановки первого типа. Опыт имитационного моделирования показал, что без указанной предварительной оценки типа перестановки с высокой вероятностью возникают коллизии и задержки в обработке информации, приводящие к зацикливанию процедуры декодирования из-за наличия обратной связи в декодере. В этой связи возникла идея разделения потоков первого и второго типов во времени. С этой точки зрения проявляется целесообразность выявлять перестановки на ранних этапах обработки данных без сортировки мягких решений символов (МРС), например, на этапе объединения жестких и мягких решений, доводя коррекцию значимых перестановок до комбинаций первого типа, минуя условия возникновения коллизий. Следовательно, отпадает необходимость в организации обратной связи в середине общего цикла при исправлении перестановки второго типа. Кроме того, нет необходимости обязательной сортировки жестких решений по показаниям МРС в порядке убывания их значений. Сохранение естественного следования нумераторов надежных символов формирует перестановку не в виде случайной последовательности нумераторов символов, а в виде их естественной возрастающей последовательности, что способствует снижению сложности вычислительного процесса за счет исключения сортировки нумераторов в порядке их убывания, что способствует рациональной организации памяти когнитивных карт.

Суть метода распространения доверия (МРД) и его особенности

В целевой функции управления для систем реального времени решающее значение играет темп обработки данных для получения итогового результата в процедуре обработки команд управления. В этом случае необходимо выявить приоритеты между задержками в получении окончательного результата. Аналитическая модель алгоритма итеративных преобразований имеет вид:

$$Ld_1 \boxplus Ld_2 \approx (-1) \times sign(Ld_1) \times sign(Ld_2) \times \min[|Ld_1|, |Ld_2|],$$

где обозначение $\boxplus$ символизирует действия по учету знаков и абсолютных значений МРС символов Ld_1 и Ld_2; при этом значение Ld_1, по сути, представляет алгебраическое сложение некоторой априорной оценки и соответствующей ей апостериорной оценки МРС на очередном шаге итерации, т. е. $Ld_1 = (d_{\text{апр}}) + (d_{\text{апос}})$ как статистически независимые информационные биты.

Важно, что для коррекции ошибок целесообразно иметь высокие значения МРС для символов проверки четности. Однако вопрос динамического изменения МРС при исправлении жестких решений символов подробно не исследовался, но принципиально положительные свойства алгоритмов МРД оказываются полезными при реализации ПД с использованием предлагаемого в работе метода локализации перестановок. Укажем главные характеристики полученных результатов.

Во-первых, если в выделенных для коррекции МРС выполняется правило четности по $mod\ 2$ для двоичных избыточных кодов, то повышение индексов МРС осуществляется за два или три цикла преобразований. В цикле, как правило, реализуются 8 арифметических операций.

Во-вторых, высокий уровень проверочных МРС однозначно позволяет исправить одиночные ошибки в информационных разрядах проверочных уравнений, при этом уровень МРС правильного разряда понижается, но ошибочный разряд меняет свой знак.

В-третьих, повысить индексы МРС проверочных разрядов невозможно без циклических сдвигов нумераторов проверочного соотношения, что отрицательно сказывается на времени обработки данных. Однако цена таких потерь в условиях коррекции данных не изучалась.

Допускается вычеркивание символа с высоким МРС, но при этом требуется коррекция показателя степени у коэффициента (-1) выражения (1) в формате $(-1)^{1+i}$, где i – число удаленных из проверочного соотношения МРС со знаком (+) [11].

Принцип двухпороговой идентификации мягких решений

Предлагается использовать второй пороговый элемент в виде усреднённой оценки МРС (λ ср) на длине кодового вектора. В этом случае существенно упрощается задача формирования текущей перестановки нумераторов символов. Пусть перестановка требует набора k разрядов и λ_i ($i \in 1, k$), значения МРС распределились в порядке убывания их абсолютных величин, т. е.

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_k|.$$

Пусть при этом, оценка переставленной матрицы $G_{\text{пер}}$, составленной из столбцов порождающей матрицы основного кода $G_{(n,k)}$, на предмет линейной зависимости показывает, что определитель такой матрицы не равен нулю. В принципе, не важно, в какой последовательности будут размещены столбцы матрицы $G_{\text{пер}}$, определитель всех матриц из подмножества k! будет равен нулю, в силу линейной независимости столбцов матрицы $G_{\text{пер}}$.

Отсюда вытекает важный технический результат: нет необходимости ранжировать показатели МРС с наиболее надежными оценками. Это означает, что все МРС, для которых справедливо $|\lambda_i| \geq |\lambda_{\text{ср}}|$,могут быть использованы для формирования ЭК. Пусть в системе используется блоковый код Хэмминга (7, 4, 3), пусть приёмник принял последовательность

$$1_6 0_3 1_2 0_7 0_5 1_3 0_6, \text{ отсюда } \lambda_{\text{ср}} = \frac{34}{7} \cong 4{,}4.$$

Выстраиваем вектор $\boldsymbol{k} = 1_6^1\ 0_7^4\ 0_5^5\ 1_6^7$, где верхний индекс указывает на нумератор столбца матрицы $G_{(7,4)}$ кода Хэмминга не в порядке возрастания МРС, а в порядке их следования в перестановке. Получаем образец перестановки нумераторов (1 4 5 7). Определитель матрицы $G_{\text{пер}}$ $\triangle \neq 0$, следовательно, возможно создание ЭК.

Таким образом, не важно, что столбцы матрицы не ранжированы по значениям МРС, но они упорядочены нумераторами столбцов, что более важно, чем ранжирование их по оценкам надежности, поскольку исключается процедура ранжирования. Это приводит к прямой экономии вычислительного ресурса. В работе декодера предусмотрена специальная когнитивная карта, в которой, например, для перестановки нумераторов (1 4 5 7) будет сразу показана проверочная часть соответствующей порождающей матрицы для формирования ЭК с простой подстановкой слева единичной матрицы для получения порождающей матрицы такого кода.

В случае, если число символов с оценками оказывается меньше, чем значение $\boldsymbol{k}$, то используется МРД для вычисления недостающих оценок.

Пусть в нашем примере приемником была зафиксирована последовательность

$$1_6\ 0_3\ 1_2\ 0_3\ 0_5\ 1_3\ 1_6, \text{ откуда } \lambda_{\text{ср}} = \frac{28}{7} = 4$$

Выстраиваем вектор $\boldsymbol{k}$=$1_6^1\ 0_5^5\ 1_6^7$, что не соответствует требованиям МРД. Декодер выделяет проверочные соотношения, соответствующие выражению $\chi_1 \oplus \chi_2 \oplus \chi_3 = h_1$, которое

определяется первым столбцом проверочной матрицы кода, и оценивает возможность его применения для увеличения индекса надёжности у символа $\chi_2 = 0_3^2$.

Выделим это уравнение $1_6^1\ 0_3^2\ 1_2^3\ 0_5^5$ и представим его в терминах МРД: $+6-3+2=-5$.

В последовательности максимальных оценок не хватает одного символа, поэтому целесообразно применить МРД к символу -3. Удаляем символ $+2$ и, чтобы не нарушать проверку чётности в уравнении (1) в значении (-1) увеличиваем степень на единицу, что приводит к положительному значению $(-1)^{1+1}=1$ этой части формулы (1).

Тогда из $+6-3=-5$ получаем:

Шаг 1	Шаг 2
$-3+0\|-5\longrightarrow +3$	$-3-5\|-5\longrightarrow +5$
$+6+0\|-5\longrightarrow -5$	$+6+3\|-5\longrightarrow -5$

Этих шагов вполне достаточно для получения откорректированной последовательности оценок. Получаем: $1_7^1\ 0_7^2\ 1_2^3\ 0_3^4\ 0_5^5\ 1_3^6\ 1_6^7$ и $\lambda_{\text{ср}} = \frac{33}{7} \cong 4$.

Выстраиваем вектор $1_7^1\ 0_7^2\ 0_5^5\ 1_6^7$. Для данной перестановки Определитель $\triangle \neq 0$. Следовательно ЭК допустимо сформировать.

Средние значения $\lambda_{\text{ср}}$ удобно использовать не только в качестве второй границы, но и в качестве критерия постановки преднамеренной помехи. В этом случае резкое снижение $\lambda_{\text{ср}}$ указывает либо на потерю синхронизации и соответственно потерю связи, либо является объективной оценкой влияния деструктивной помехи.

Заключение

Введение двухуровневого порога образования МРС и их внедрения в порядке следования нумераторов позиций существенно снижает сложность реализации процесса формирования порождающей матрицы ЭК и, главное, поиск подобной матрицы в когнитивной карте декодера.

Обоснованный ввод среднего значения МРС на длине кодового вектора позволяет управлять объектом за счёт своевременной реакции на вероятное применение преднамеренных помех.

Необходимо осуществить имитационное моделирование для получения обоснованных статистических материалов.

Литература

1. Гладких А.А. Перестановочное декодирование как инструмент повышения энергетической эффективности систем обмена данными // Электросвязь. 2017. № 8. С. 52–56.

2. Свойства циклических структур в системе перестановочного декодирования избыточных кодов / Бабанов Н.Ю., Гладких А.А., Наместников С.М., Шахтанов С.В. // Автоматизация процессов управления. 2020. № 2 (60). С. 101-108.

3. Оценка критериев эффективности защиты данных от ошибок на базе перестановочного декодирования / А.В. Новоселов, С.В. Шахтанов, А.С.А. Аль-Месри, М.Ю. Толикина // Автоматизация процессов управления. 2022. № 3 (69). С. 27–34.

4. Перестановочное декодирование с системой адаптированных альтернативных решений / А.А. Гладких, А.А. Овинников, Н.А. Пчелин, А.А. Брынза // Цифровая обработка сигналов. 2023. № 4. С. 73–78.

METHOD FOR GENERATING PERMUTATIONS AND CORRECTING DEGENERATE MATRICES OF A PERMUTATION DECODER

Potapova S.E., Ulasyuk T.G., Ulyanovsk Institute of Civil Aviation

The report discusses the method of permutation decoding of short block error-correcting codes, which provides acceptable data reliability in conditions of sufficiently low signal-to-noise ratios. This is achieved by increasing the multiplicity of errors corrected by the code, which is typically equal to the number of redundant symbols. However, this method has a disadvantage when applied to binary redundant codes, as it introduces a high probability of permutations of the numerators of the received code combination, resulting in a degenerate permuted matrix that cannot be corrected to obtain an equivalent code. To achieve a positive result, it is proposed to use the trust propagation method, which allows for the required correction of exactly those numerators of the characters to be permuted, which allow for the most efficient obtaining of an equivalent code.

МОДЕЛИРОВАНИЕ СИСТЕМЫ РАЗБИЕНИЯ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ НА КЛАСТЕРЫ

асп. Ваняев Д.А.

Нижегородский государственный инженерно-экономический университет

В работе представлен алгоритм организации создания кластера робототехнических устройств. Проведено моделирование для эквивалентных кодов на основе кода Хэмминга (15, 11) с формированием уникальных невырожденных матриц по итогам перестановок. Было получено всего 1365 сочетаний из которых 839 являются уникальными и невырожденными. Проведено распределение матриц на кластеры по принципу минимального евклидового расстояния между их векторами. В кластерах выделены центроиды. Предложен способ практического применения повышения безопасности передачи данных, переключением на другую матрицу с использованием псевдослучайного генератора чисел.

Внедрение в современную жизнь групп робототехнических устройств, исполняющих единую задачу, в составе: беспилотных авиационных систем, наземных мобильных и статичных манипуляторов, подводных и подземных устройств, выявило проблему организации связи между ними и наземной станцией управления (НСУ) [1, 2]. Одно из очевидных и используемых решений является разделение всей сети устройств на кластеры. Каждый кластер имеет собственную задачу и устройства в кластере работают совместно для достижения единой целевой функции.

Для деления на кластеры можно использовать различные методы, которые разделяются на большие группы: мягкие и жёсткие. При жёсткой кластеризации устройство может входить только в один кластер, в мягком устройства могут принадлежать нескольким кластерам [3]. Для выполнения кластеризации необходимо выделить определённый алгоритм, по которому устройства будут распределены в группы. Данный алгоритм был разработан [4] и представлен ниже:

1. Каждому устройству заранее назначается уникальный адрес в общей сети.
2. Для всех устройств задаются исходные параметры, включая тип платформы, выполняемую задачу полезной нагрузки и текущий уровень износа.
3. Оператор на НСУ самостоятельно задаёт целевую функцию, определяющую критерии выбора устройств.
4. Алгоритм анализирует параметры всех устройств, вычисляет их соответствие заданной цели и присваивает каждому значение пригодности в диапазоне от 0 до 1. На основе этих значений выполняется распределение устройств по кластерам.
5. Из сформированного списка алгоритм выбирает наиболее подходящие устройства, объединяет их в кластер и назначает ему ключевую комбинацию. Для обеспечения корректного обмена данными каждому дрону внутри кластера присваивается индивидуальный номер. В процессе обмена сообщениями НСУ и устройства подтверждают получение и установку ключевой комбинации.
6. После формирования кластера и передачи кодовых слов начинается выполнение задачи, за ходом которой оператор наблюдает через НСУ.
7. По завершении задачи кластер либо расформировывается, либо переводится в режим ожидания следующего задания.

Данный алгоритм решает проблему распределения адресов внутри кластера, но не рассматривает способы создания адресов и защиты передаваемой информации по каналу управления. При выполнении целевой задачи система должная находиться в состоянии гомеостаза, то есть способности открытой системы сохранять постоянство своего внутреннего состояния посредством скоординированных реакций, направленных на поддержание динамического равновесия.

За основу системы разделения кодовых слов, которые используют устройства и НСУ, использовались эквивалентные коды применяемы к системе обработки блоковых систематических кодов [5]. В подобных системах на приемной стороне реализуется метод мягкого декодирования кодовых последовательностей, поступающих из канала связи, подверженного воздействию помех. На основе оценки достоверности принятых символов

выполняется их перестановка таким образом, чтобы наиболее надёжные элементы занимали позиции информационных разрядов.

Параллельно с этим, в соответствии с выбранной перестановкой формируется порождающая матрица соответствующего эквивалентного кода. С её помощью наиболее достоверные информационные символы повторно кодируются уже в условиях отсутствия канальных искажений. Полученный кодовый вектор рассматривается как эталонный, обладающий большей надёжностью по сравнению с принятым из канала. Сопоставление эталонной и принятой кодовых комбинаций позволяет выявить локаторы ошибок и оперативно выполнить их исправление. При этом для каждой принятой кодовой комбинации формируется индивидуальный ЭК, определяемый конкретной перестановкой символов [6].

Построение модели расчёта матриц и кодовых слов для кода (15, 11)

Реализация способа разбиение пространства перестановок на кластеры было рассмотрено в работе на примере кода Хэмминга (7, 4) [5]. По итогу выполнения расчётов было выявлено, что при таком количестве информационных и избыточных разрядов общее количество уникальных перестановок, которые возможно использовать при создании кодовых слов для канала управления равны 24. Этого недостаточно для полноценной самоорганизующейся сети. Было решено увеличить количество информационных и избыточных разрядов до кода (15, 11). Использован порождающий многочлен (1).

$$g(x) = x^4 + x + 1, \tag{1}$$

Получена порождающая матрица, приведённая к систематическому виду (2).

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \tag{2}$$

Далее необходимо выполнить ряд перестановок столбцов для получения новых матриц и соответствующих им кодовых слов. Для этого рассчитаем количество вырожденных матриц, в которых определитель равен нулю, так как в последствии их невозможно будет использовать.

По результатам моделирования были получены следующие данные: количество сочетаний $C_7^4 = 1365$, из которых вырожденных 525 и одна матрица повторяется в ходе перестановок. Итого пригодных для работы 839 матриц. Первые и последний представлены в таблице 1. Проверочные матрицы представлены в систематической форме, которая используется микропроцессором декодера для вычисления синдрома. Процесс заключается в умножении принятого кодового вектора на соответствующую матрицу. Нулевой результат вычисления (нулевой вектор синдрома) свидетельствует об отсутствии обнаруженных ошибок в принятой комбинации избыточного кода.

Пригодные матрицы были распределены на 15 кластеров. Для работы преобразовываем матрицы в вектор из 44 бит, результат каждая матрица – точка в 44-мерном пространстве. Алгоритм группирует матрицы по минимальному евклидовому расстоянию между их векторами. В каждом кластере выбрана главная матрица – аналог центроида, элемента являющегося наиболее близким к среднему значению всех матриц в кластере. В каждом кластере находится от 30 до 68 элементов включая центроид, соответственно данные значения равны максимальному количеству устройств для кластера. Данные по некоторым матрицам и центроидам представлены в таблице 2.

Таблица 1. Уникальные невырожденные матрицы после перестановок

Перестановка	[4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15]	[4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15]	[3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 15]	[3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14]	[3, 4, 5, 6, 7, 8, 9, 12, 13, 14, 15]
Матрица	0 0 1 1 1 1 0 0 0 1 1 0 1 1 1 1 1 0 1 0 0 1 1 1 1 1 1 0 0 1 0 1 1 1 0 1 1 0 0 1 1 0 1 1	1 1 0 1 1 1 0 0 0 1 1 0 1 1 1 1 1 0 1 0 1 0 0 1 1 1 1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 0 1	1 0 1 0 0 0 1 1 1 1 0 0 1 1 1 0 1 0 0 1 1 1 1 1 0 1 1 1 0 1 1 0 1 1 0 1 0 1 0 1 1 0 1 1	1 0 1 0 1 0 0 1 1 1 0 0 1 1 1 0 0 0 1 1 0 1 0 1 1 1 0 1 0 1 1 0 0 1 1 1 1 1 1 1 1 0 1 1	1 1 0 1 0 1 1 0 1 1 0 0 1 0 0 1 1 0 1 1 1 0 1 0 0 1 0 1 1 1 1 1 0 1 1 1 0 0 1 1 1 1 1 0

Таблица 2. Кластеры и количество элементов в них

№ кластера	1	2	3	4	5	6	…	15
Центроид	1 1 1 1 1 0 1 1 0 0 1 1 1 1 0 1 0 1 1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 0 1 1 0 1 1 1 1 1	1 1 1 0 1 1 1 0 1 0 1 1 0 1 1 1 1 1 0 1 0 1 1 1 0 1 0 1 1 1 1 0 1 1 1 1 1 1 0 1 0 0 1 1	1 1 1 1 1 1 1 0 1 1 1 0 0 1 1 1 1 1 0 1 0 0 1 1 1 0 1 1 1 1 1 0 1 0 1 1 0 1 0 1 1 1 1 1	0 1 1 1 1 0 1 1 0 1 0 1 1 1 0 0 1 1 1 1 1 1 0 1 1 0 1 1 1 1 1 1 1 1 1 1 1 0 1 0 0 1 1 1	0 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1 0 1 0 1 0 1 1 0 1 1 1 1 0 0 1	1 0 1 1 1 0 1 1 0 1 1 1 1 0 1 1 1 1 0 1 1 1 0 1 1 1 0 0 1 1 1 1 1 1 0 1 1 0 0 1 1 1 1 0	… … … … … … … … … … …	0 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 0 0 0 1 1 1 1 0 0 1 0 0 1 1 1 1 0 1 1 0 1
Количество элементов в кластере	63	63	64	52	64	53	…	45

При распределении кодовых слов устройствам, каждому робототехничскому средству присваивается ключевая комбинация общее для всех кластеров полученное из порождающей матрицы, ключевая комбинация индивидуально для его кластера и его собственная ключевая комбинация. При отправке сообщения в эфир соответствующим кодовым словом, получить его могут все, а раскодировать только те устройства, которым оно предназначается.

При отправке сообщения от наземной станции с номерами столбцов матрицы {2,3,4,5,6,7,8,9,10,12} к устройству с номерами столбцов матрицы {4,5,6,8,9,10,11,12,13,14,15}, декодер в наземной станции произведёт перестановки для совпадения с матрицей получателя, произведёт кодирование, передатчик отправит команду устройству. Аналогичным образом происходит процесс передачи команд для других устройств и всему кластеру в целом.

Практическое применение

Интенсификация разработок в сфере робототехнических систем способствует их интеграции в системы обеспечения транспортной безопасности и многих других сфер деятельности. Гетерогенная архитектура коллективного взаимодействия робототехнических комплексов позволяет делегировать им функции мониторинга и контроля, традиционно закрепленные за человеческим ресурсом. При этом топология защищаемых объектов варьируется от распределенных магистральных структур до компактных узловых центров

Важной задачей в ходе передачи данных является безопасность. В условиях, когда количество матриц существенно превышает количество устройств появляется возможность изменять данные перестановки, что повысит безопасность передачи команд управления. Для реализации такой системы необходимо внести генератор псевдослучайной частоты на кодер и декодер, провести синхронизацию между устройствами для корректной передачи данных и смены матрицы.

Заключение

В работе рассмотрено допустимое множество ЭК для кода Хэмминга (15, 11). Выполнены перестановки и рассчитано количество уникальных невырожденных матриц, количество которых равно 839. Проведено построение 15 кластеров с группировкой вектором матриц по минимальному евклидовому расстоянию и выделением центроида, для распределения их каждому устройству. Предложен вариант повышения безопасности передачи команд управление, изменяя матрицу с использование псевдослучайного генератора чисел. Так же предложен алгоритм распределения адресов кластера, где процесс объединения устройств в кластеры базируется на анализе функционального назначения группы, технических характеристик роботов, а также их пространственного расположения относительно друг друга и наземной станции управления.

Литература

1. Агеев С.А. Метод и алгоритмы подсистемы оперативной поддержки принятия решений для управления техническим состоянием судового комплекса связи / С.А. Агеев, Н.С. Агеева, А.А. Гладких, В.В. Каретников // Автоматизация процессов управления. 2024. № 2 (76). С. 32–42.

2. Башилов, А. М. Беспилотные технологии в системах аграрного производства / А. М. Башилов, В. А. Королев // Вестник НГИЭИ. – 2023. – № 5(144). – С. 43-56. – DOI 10.24412/2227-9407-2023-5-43-56. – EDN EEDXDV.

3. Егоров, А. В. Особенности методов кластеризации данных / А. В. Егоров, Н. И. Куприянова // Известия ЮФУ. Технические науки. – 2011. – № 11(124). – С. 174-178. – EDN OJCYUZ.

4. Ваняев, Д. А. Алгоритм автоматического формирования кластеров робототехнических устройств для организации связи / Д. А. Ваняев, А. А. Гладких // Проблемы техники и технологии телекоммуникаций. Оптические технологии в телекоммуникациях : материалы XXVII и XXIII Международных научно-технических конференций, Уфа, 05–07 ноября 2025 года. – Уфа: Уфимский университет науки и технологий, 2025. – С. 99-100. – EDN OVBXJC.Гладких А.А. Перестановочное декодирование как инструмент повышения энергетической эффективности систем обмена данными // Электросвязь. 2017. № 8. С. 52–56.

5. Гладких А.А. Архитектура коллективного управления с использованием свойств эквивалетности блоковых избыточных кодов / А. А. Гладких, Т. Г. Уласюк, М. Ю. Толикина [и др.] // Автоматизация процессов управления. – 2025. – № 4(82). – С. 4-11. – DOI 10.35752/1991-2927_2025_4_82_4. – EDN LXDJIG.

6. Гладких А.А. Перестановочное декодирование как инструмент повышения энергетической эффективности систем обмена данными // Электросвязь. 2017. № 8. С. 52–56.

MODELING OF ROBOTICS SYSTEMS CLUSTERING SYSTEM

Vanyaev D.A.

Nizhny Novgorod State University of Engineering and Economics

The paper presents an algorithm for organizing the creation of a cluster of robotic devices. The paper conducted modeling for equivalent codes based on the Hamming code (15, 11) with the formation of unique non-degenerate matrices based on the results of permutations. A total of 1365 combinations were obtained, of which 839 are unique and non-degenerate. The matrices were divided into clusters based on the principle of the minimum Euclidean distance between their vectors. The centroids were identified in the clusters. A method for improving data transmission security by switching to a different matrix using a pseudo-random number generator has been proposed.

АЛГОРИТМ ОЦЕНКИ ЧАСТОТЫ ИСТОЧНИКА РАДИОИЗЛУЧЕНИЯ В СИСТЕМАХ РАДИОМОНИТРОИНГА

проф. Сучков В.Б., доц. Глазков В.В., инженер Силин С.И.

МГТУ им. Н.Э. Баумана

Рассмотрены принципы алгоритма анализа частоты радиоизлучающего средства в системе многопозиционного пассивного радиолокатора. Рассмотрены общие принципы работы пассивного радиолокатора в части задачи геолокации источника излучения. Описаны проблемы, связанные с работой при большом количестве источников излучения. Предложены методы упрощения задачи геолокации. Предложен алгоритм оценки частоты принятого сигнала для решения задачи упрощения геолокации. Приведены результаты моделирования предложенного алгоритма.

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации по государственному заданию № FSFN-2025-0014.

Введение

Одним из назначений систем пассивной радиолокации [1-5] является определение наличия в относительной близости посторонних источников радиоизлучения, определение направления на них (пеленгация), распознавания их типов и определение их координат в локальной или глобальной (географической) системе координат. Задача геолокации обычно решается геодезическим способом – в нескольких удаленных точках располагаются приемники пассивного радиолокатора, после чего на каждом из них определяется направление на источник излучения. После этого линии, соответствующие пеленгам каждого из источников [6-8], наносятся на карту, и пересечение показывает положение источника радиоизлучения. Зачастую работа радиолокатора происходит в условиях, когда таких источников в поле зрения и принимаемой полосе больше одного.

На рисунке 1 проиллюстрирована работа с несколькими источниками. Два приемника принимают сигналы двух объектов, различимых по пространству, работающих в один промежуток времени и имеющие достаточно близкие параметры излучаемого сигнала. Каждый из приемников принимает два сигнала и оценивает направление, с которого сигналы пришли.

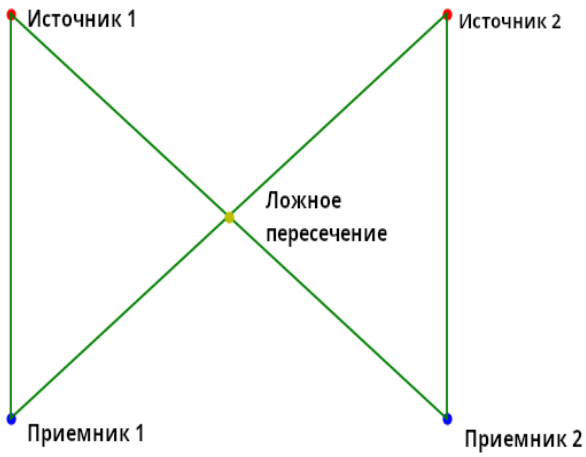


Рисунок 1 - Схема пассивной радиолокации нескольких источников излучения в многопозиционной системе

При построении линий положения источников на карте возникает три пересечения – два пересечения в точках, где источник есть, и одно пересечение «ложно» - в точке, где источника нет. При увеличении числа источников и числа приемников количество ложных пересечений возрастает. Например, при трех источниках будет три ложных пересечения, при четырех – шесть, и так далее.

Для избавления от ложных пересечений необходим дополнительный анализ, связанный с накоплением пересечений, их кластеризацией и анализом параметров полученных кластеров для определения достоверности соответствия кластера реальному источнику излучения. Подобные алгоритмы требовательны к вычислительным ресурсам и в связи с необходимостью накопления, увеличивают задержку в принятии решения относительно наличия или отсутствия источника излучения в заданной точке. Другой подход связан с группировкой сигналов в зависимости от их параметров. Предполагая, что каждый (даже однотипный) источник обладает уникальными чертами, ведущими к расхождению между параметрами принятых сигналов, можно попытаться увеличить точность оценки параметров сигналов для выделения группы принятых сигналов, заведомо соответствующих одному источнику и свести задачу совместного поиска нескольких источников излучения к соответствующему числу отдельных задач поиска одного источника, тем самым устранив проблему ложных пересечений, и снизив время отклика системы.

Современный метод реализации таких систем в реальном времени предполагает использование модулей цифровой обработки на базе ПЛИС или заказных микросхем. При расположении модулей обработки на бортовой системе разработчики сталкиваются с серьезным ограничением ресурсов, не позволяющих производить сложные вычисления. Таким образом важной задачей является разработка алгоритмов оценки параметров для подобных систем реального времени. В текущем докладе предложен метод оценки частотных параметров.

Алгоритм оценки частоты

Предполагая, что полезный сигнал представлен в виде выборки из смеси аддитивного белого гауссовского шума (АБГШ) и комплексного гармонического сигнала от источника радиоизлучения, плотность распределения вероятностей комплексного сигнала можно записать в виде:

$$W(Z) = \prod_{i=1}^{N} \frac{1}{2\pi\sigma} exp\left(-\frac{\left|Z - Ae^{j(\omega i+\varphi_0)}\right|^2}{4\sigma^4}\right), \quad (1)$$

где A – амплитуда сигнала, N – число отсчетов в выборке, Z – выборка комплексного сигнала φ_0 – начальная фаза, ω – относительная круговая частота/

Оценка частоты по совокупности отсчетов (например, с помощью БПФ) имеет высокую вычислительную сложность для приемлемой точности, в связи с чем предлагается вычислять оценку частоту через оценку мгновенной фазы сигнала. Тогда плотность распределения вероятности фазы одного отсчета будет выглядеть следующим образом:

$$W(\varphi_i) = \frac{1}{2\pi} e^{-\frac{q^2}{2}} + \frac{q}{\sqrt{2\pi}} cos(\varphi_i - \overline{\varphi_i}) e^{-\frac{q^2}{2} sin^2(\varphi_i - \overline{\varphi_i})} \frac{1+\mathrm{erf}\left(q\frac{cos(\varphi_i - \overline{\varphi_i})}{\sqrt{2}}\right)}{2}, \quad (2)$$

где q – отношение сигнал-шум (ОСШ) $q = \frac{A}{\sigma}$, φ_i – фаза i-го отсчета $\overline{\varphi_i}$ – среднее значение фазы i-го отсчета.

На ОСШ, при которых возможно обнаружение в подобных системах функция ПРВ хорошо аппроксимируется гауссовской плотностью распределения:

$$W(\varphi_i) = \frac{q}{\sqrt{2\pi}} e^{-\frac{q^2(\varphi_i - \overline{\varphi_i})^2}{2}}. \quad (3)$$

При значениях относительной круговой частоты около 0.25 фаза, полученная через взятие аргумента от отсчета приводит к появлению неоднозначности при переходе через 2π и сложностям с заданием ожидаемых значений в модели, что затрудняет анализ. В связи с этим предлагается перейти к определению разности набега фаз в двух соседних отсчетах через вычисление аргумента

$$\Delta\varphi_i = \varphi_i - \varphi_{i-1} = \overline{\varphi_i} - \overline{\varphi_{i-1}} + n_i - n_{i-1} = arg(z_i z^*{}_{i-1}), \quad (4)$$

где n_i- шум фазы.

С учетом преобразования фазы приближенную совместную плотность распределения фазы можно записать в виде:

$$W(\varphi_1 \dots \varphi_N) = \sum_{i=1}^{N} \frac{q}{\sqrt{2\pi}} exp\left(-\frac{q^2(\varphi_i - \overline{\varphi_i})^2}{2}\right), \quad (5)$$

$$\varphi_i = \sum_{j=2}^{i} \Delta\varphi_j = \sum_{j=2}^{i} \overline{\varphi_i} + n_i - n_1, \quad (6)$$

где $\overline{\varphi_i} = \omega i + n_1$.

Тогда с учетом выражений (1-6) можно найти оценку частоты через решение уравнения:

$$AX = B, \tag{7}$$

$$A = \begin{bmatrix} \sum_i i^2 & \sum_i i \\ \sum_i i & N \end{bmatrix}, \tag{8}$$

$$X = \begin{bmatrix} \omega \\ n_1 \end{bmatrix}, \tag{9}$$

$$B = \begin{bmatrix} \sum_i i\varphi_i \\ \sum_i i \end{bmatrix}. \tag{10}$$

На рисунках 2-3 представлены результаты численного моделирования среднеквадратической ошибки относительной частоты источника радиоизлучения в зависимости от ОСШ по формулам (7-10).

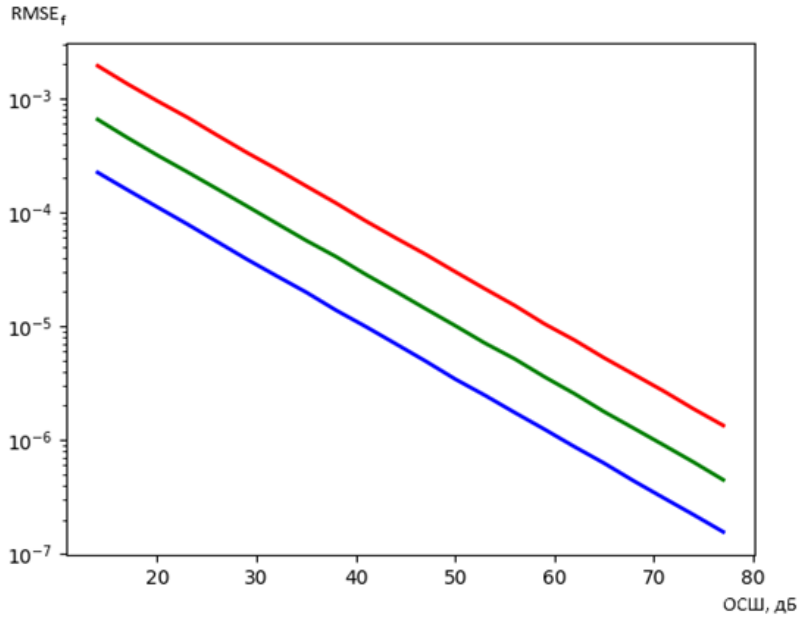


Рисунок 2 - Зависимость среднеквадратической ошибки относительной частоты от ОСШ (красная кривая –размер выборки 16, зеленая кривая –размер выборки 32, синяя кривая - размер выборки 64)

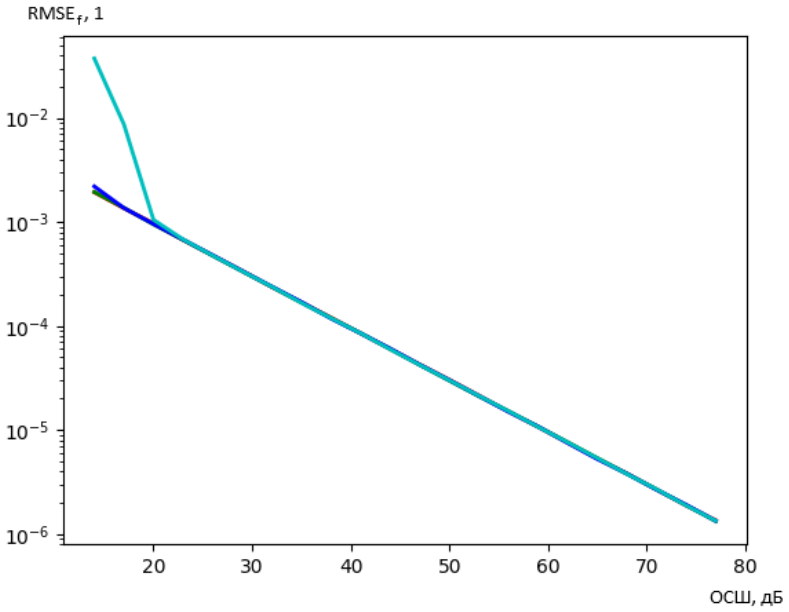


Рисунок 3 - Зависимость среднеквадратической ошибки относительной частоты от ОСШ (черная кривая – относительная частота 0.1, синяя кривая – относительная частота 0.2, бирюзовая кривая – относительная частота 0.3)

Из графиков на рисунке 3 видно, что при ОСШ более 20 дБ частота сигнала не оказывает влияния на точность измерений. Для увеличения точности оценки возможно использовать M приемных каналов в каждой приемной системе, уменьшив дисперсию оценки в M раз (среднеквадратическую ошибку в корень из M раз).

Сравним вычислительную сложность и потребные вычислительные ресурсы разработанного алгоритма с алгоритмом на основе БПФ. Алгоритм БПФ можно представить (рисунок 4) в виде послойной комбинации ячеек, состоящих из комплексного умножителя (4х вещественных) и комплексного сумматора (2х вещественных). Для вычисления в реальном времени требуется NlogN таких ячеек, где N – число анализируемых отсчетов, либо logN ячеек и высокопроизводительная (2*logN*f *W, б/с) память для хранения NlogN отсчетов, где f – частота тактового сигнала, W – ширина одного спектрального отсчета в битах. При больших значениях N реализация подобного алгоритма в истинном реальном времени затруднительна

Схема разработанного в данной работе алгоритма представлена на рисунке 5 в виде конвейера вычисления фазы отсчета и пяти итерационных аккумуляторов, считающих коэффициенты матриц A и B. При этом количество элементов алгоритма не зависит от N.

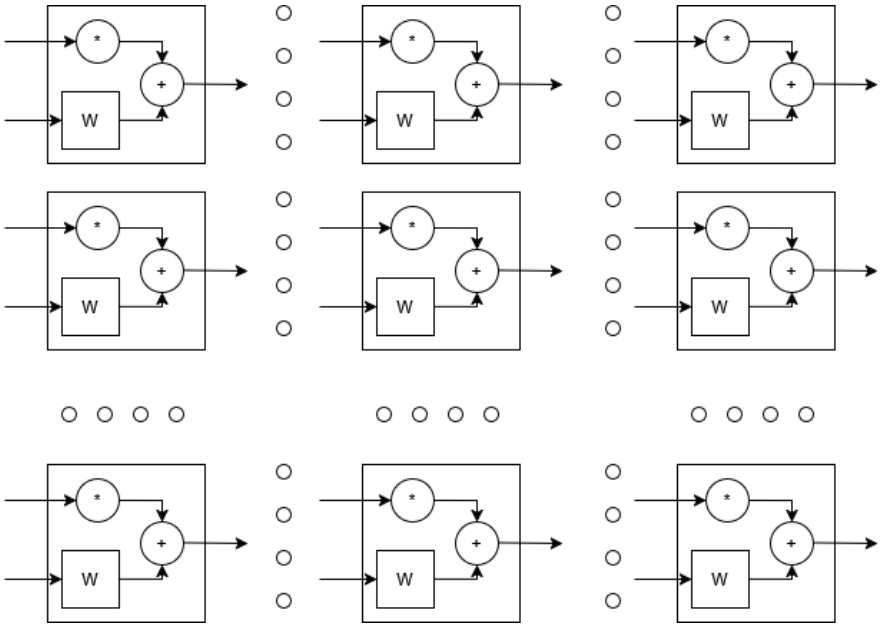


Рисунок 4 – Схема вычисления БПФ

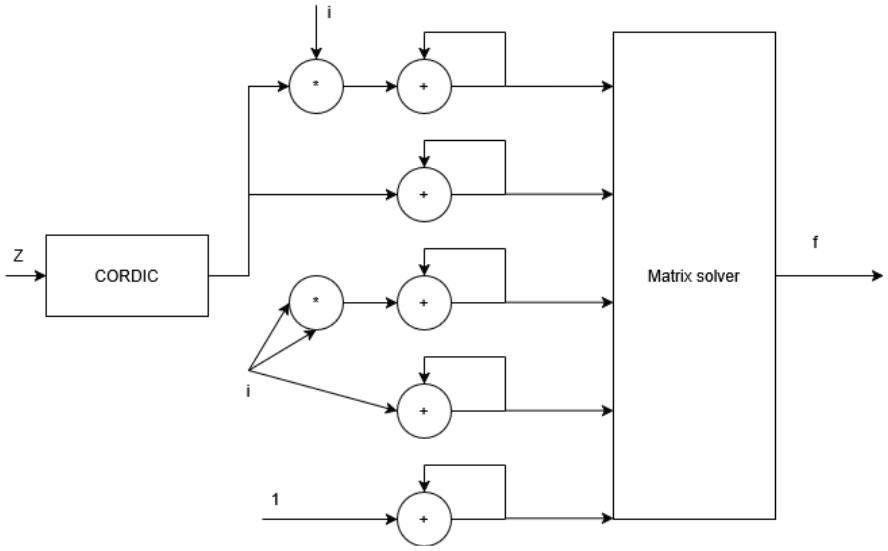


Рисунок 5 – Схема вычислений дискретных отсчетов на основе алгоритма в виде конвейера вычисления фазы отсчета и пяти итерационных аккумуляторов

Таким образом затраты вычислительных ресурсов на выполнение в реальном времени ниже, чем у БПФ при выполнении предположения о том, что входные сигналы разделены и не накладываются друг на друга при оценке частоты каждого из них.

Заключение

В работе подтверждена эффективность предложенного алгоритма анализа частоты радиоизлучающего средства в системе многопозиционного пассивного радиолокатора. Предложен алгоритм оценки частоты в бортовых системах реального времени, предназначенных для пассивной радиолокации для случая ограниченных вычислительных ресурсов, позволяющий повысить скорость геолокации. Проведен расчет точности оценки частоты, и предложена методика выбора архитектуры системы для достижения желаемой точности.

Литература

1. Бузин А.Л., Парфенов Г.Г. Особенности использования средств радиотехнической разведки как источников (средств) целеуказания для радиоэлектронной борьбы // Радиоэлектронная борьба. 2022. Т. 3. № 2. С. 17−22.

2. Ву Б.Х. Способы и алгоритмы разрешения и оценивания параметров нескольких сигналов в радиосистемах Автореф. дис. канд. техн. наук. Рязань: Рязанский гос. радиотехнич. ун-т. 2024. 20 с.

3. Горбунов Ю.Н. Рандомизированная обработка сигналов в радиолокации и связи. Монография. М.: Изд-во «LAP LAMBERT Academic Publishing». 2015. 150 с.

4. Алдаков А.Х. Способ аналитического расчета пространственного положения источника сигнала на основе разностно дальномерного метода // Радиотехника. 2025. Т. 89. № 10. С. 135−139.

5. Горбунов Ю.Н., Акопян Г.Л., Климова С.А. Цифровая обработка пространственных и временны́х сигналов с применением параметров оконных весовых функций для систем радиотехнического мониторинга // Радиотехника. 2025. Т. 89. № 10. С. 146−154.

6. Бирюков И.Д. Применение совместных алгоритмов обработки радиосигналов в системах радиотехнического наблюдения // Радиотехника. 2025. Т. 89. № 10. С. 122−134.

7. Созинов П.А., Андреев Г.И., Тихомиров В.А., Замарин М.Е. Перспективы развития средств радиотехнического мониторинга в радиоэлектронной борьбе // Радиотехника. 2025. Т. 89. № 10. С. 61−71.

8. Силин С. И. Алгоритмы обнаружения сигналов пассивной радиолокации и оценка их фазочастотных параметров: диссертация на соискание ученой степени кандидата технических наук / Силин Сергей Игоревич. – Москва, 2012. – 148 с. – EDN QFXMLR.

ALGORITHM FOR ESTIMATION OF THE FREQUENCY OF RADIO SOURCE IN RADIO MONITORING SYSTEMS

Suchkov V.B., Glazkov V.V., Silin S.I.

Bauman Moscow State Technical University

The principles of the algorithm for analyzing the frequency of a radio-emitting device in a multi-position passive radar system are discussed. The general operating principles of passive radar are considered in relation to the task of geolocating the radiation source. The problems associated with working with a large number of radiation sources are described. Methods for simplifying the geolocation task are proposed. An algorithm for estimating the frequency of the received signal is proposed to solve the geolocation problem. Simulation results of the proposed algorithm are presented.

ИСПОЛЬЗОВАНИЕ МЕТОДОВ ТЕСТОВЫХ ИСПЫТАНИЙ ДЛЯ МАТЕМАТИЧЕСКОЙ РЕКОНСТРУКЦИИ МНОГОМЕРНЫХ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

с.н.с., к.т.н. Новиков-Бородин А.В.
Институт ядерных исследований РАН, 117312 Москва, Россия

Аннотация: Рассматриваются возможности применения математических методов тестовых испытаний для реконструкции многомерных сигналов, в том числе изображений, искаженных при обработке линейными инвариантными системами. Линейные инвариантные системы охватывают широкий класс систем обработки, описываемых многомерными уравнениями типа свертки, и являются расширением понятия линейных стационарных систем на пространственные координаты. Методы тестовых испытаний позволяют осуществить реконструкцию сигналов без определения передаточных функций систем обработки, относящихся к классу обобщенных функций, вместо которых используются тестовые сигналы из класса основных функций. Использование тестовых сигналов, значительно упрощает процедуру реконструкции и делает ее более точной и устойчивой. Операторный подход, основанный на уравнении многомерной свертки, повышает скорость численных вычислений. Для решения некорректно поставленных и плохо обусловленных задач используется техника регуляризации, позволяющая эффективно восстанавливать реальные недетерминированные сигналы с высоким уровнем шумов и неопределенностей. Приводятся примеры использования методов, процедура реконструкции сравнивается с существующими подходами. Предлагаемые методики позволяет расширить область применения математических методов реконструкции, повысить качество обработки сигналов и изображений без использования сложных и дорогостоящих аппаратных методов модификации оборудования.
Программные коды и наборы данных доступны по адресу: https://github.com/novikov-borodin/data-rec.
Ключевые слова: математическая обработка сигналов и изображений; линейные инвариантные системы; тестовые испытания систем обработки; многомерное уравнение типа свертки; некорректно поставленные и плохо обусловленные задачи; деконволюция; техника регуляризации.

Введение

Реальные системы обработки несовершенны, что приводит к системным искажениям и появлению случайных шумов в обрабатываемых ими сигналах. Аппаратные методы уменьшения искажений связаны с решением сложных научно-технических задач и созданием дорогостоящего оборудования, а альтернативные математические методы сталкиваются с серьезными трудностями при решении некорректно поставленных и плохо обусловленных задач реконструкции, как правило, встающих на практике (см., например, [1]).

Линейные инвариантные (ЛИ) системы, рассматриваемые в данной работе, удовлетворяют многомерному уравнению типа свертки (см. раздел 1) и являются инвариантными не только по временным, как стационарные, но и по пространственным координатам. ЛИ-системы описываются передаточными функциями, характеризующими системные искажения при измерении, приеме, передаче или другой обработке сигналов. В неслепых методах математической реконструкции сигналов используются передаточные функции систем обработки, но эти функции принадлежат к классу обобщенных, и их определение является достаточно сложной задачей, что резко ограничивает точность неслепых методов и область их применения некоторыми частными случаями известных видов искажений и/или шумов (см., например, [2-4]). Методы слепого восстановления (см., например, [5]) позволяют избежать проблем с определением передаточных функций, но они основаны на априорных знаниях или предположениях об исходном или желаемом виде реконструируемых данных. Реальные сигналы, искажения и шумы, вносимые системами обработки, очень разнообразны, что сильно ограничивает область применения слепых методов. Обойтись без прямого определения передаточных функций позволяют нейронные сети (см., например, [6]), но их обучение происходит на большом количестве примеров, что дает усредненные результаты и во многих случаях не обеспечивает высокой точности реконструкции. По общим оценкам, приведенным в пленарном докладе [7], существующие математические методы, включая нейронные сети, позволяют повысить эффективные параметры измерительных систем (например, эффективную полосу пропускания) в среднем в $1{,}5 - 2$ раза.

В данной работе рассматриваются особенности и примеры практического применения метода тестовых испытаний (ТИ) [8,9], совмещающего в себе отдельные черты перечисленных методов, определяющие его преимущества. Формально, ТИ-метод не использует передаточные функции

систем обработки и не является неслепым, но ТИ-уравнение, лежащее в его основе, устанавливает точное соответствие между сигналами. Использование тестовых сигналов вместо передаточных функций можно рассматривать как процесс обучения нейросети при поиске обратного преобразования. В отличие от передаточных функций, тестовые сигналы принадлежат к классу основных функций, поэтому их определение, как численными, так и экспериментальными методами, не вызывает больших проблем, что определяет простоту, универсальность и точность предлагаемого подхода. При отсутствии передаточных функций из класса обобщенных появляется возможность эффективного использования существующих методов решения некорректно поставленных и плохо обусловленных задач реконструкции, таких как методы регуляризации (см., например, [12]), используемые в данной работе, или итерационные методы, вейвлет анализ и др. Эти особенности обеспечивают универсальность предлагаемого подхода, точность восстановления и устойчивость алгоритма даже при высоком уровне шумов и неопределенностей в данных. В работе приводятся примеры реконструкции сигналов, проводится сравнение предлагаемой методики с существующими подходами.

1. Линейные инвариантные системы

Функциональную систему обработки одномерных или многомерных сигналов, включая изображения, будем называть линейной инвариантной (ЛИ), если в области $x = (x_1, x_2, \dots) \in D$ она характеризуется передаточной функцией $u(x)$ и ее реакция $w(x)$ (выходной сигнал) на внешнее воздействие $v(x)$ (входной сигнал) описывается многомерным уравнением типа свертки, которое будем называть уравнением ЛИ-систем, или, кратко, ЛИ-уравнением:

$$w(x) - \{\eta(x)\} = (u * v)(x) = \int_D u(x - \xi)v(\xi)\mathrm{d}\xi. \qquad (1)$$

Неопределенность ЛИ-уравнения: $\eta(x) = (\Delta w - u * \Delta v - \Delta u * v + \Delta u * \Delta v)(x)$ можно определить из уравнения свертки: $w_0(x) = (u_0 * v_0)(x)$ детерминированных функций без неопределенностей: $w_0(x) = (w - \Delta w)(x)$, $u_0(x) = (u - \Delta u)(x)$ и $v_0(x) = (v - \Delta v)(x)$.

Вводя оператор свертки $\mathcal{F}$ с порождающей функцией $f(x)$: $\mathcal{F}g(x) = (f * g)(x)$, ЛИ-уравнение можно представить в операторной форме: $\mathcal{U}v(x) = w(x) - \{\eta(x)\}$, а используя теорему Бореля о свертке, – в спектральной форме: $U(\omega)V(\omega) = W(\omega) - \{N(\omega)\}$, где $F(\omega) = FT[\mathcal{F}] = FT[f(x)]$ – спектр (преобразование Фурье) функции $f(x)$ и оператора $\mathcal{F}$, и $N(\omega) = \Delta W(\omega) - U(\omega)\Delta V(\omega) - \Delta U(\omega)V(\omega) - \Delta U(\omega)\Delta V(\omega)$.

ЛИ-уравнение в дискретном виде представляет собой частичную сумму интеграла (1):

$$w_{i_1..i_M} - \{\eta_{i_1..i_M}\} = \sum_{j_1=1}^{m_1} \cdots \sum_{j_M=1}^{m_M} u_{j_1..j_M}\, v_{i_1-j_1..i_M-j_M} \Delta_{j_1..j_M}, \qquad (2)$$

где $w = \{w_{i_1..i_M}\}$, $\eta = \{\eta_{i_1..i_M}\}$, $u = \{u_{i_1..i_M}\}$ и $v = \{v_{i_1..i_M}\}$ – M-мерные массивы, соответствующие функциям $w(x)$, $\eta(x)$, $u(x)$ и $v(x)$, $\Delta_{j_1..j_M}$ – элемент объема, представляет собой частичную сумма интеграла (1) на сетке $\{d_{i_1\dots i_M}\}$, $i_k = 1 \dots m_k, k = 1 \dots M$, разбивающей прямоугольную область $D_0 \supseteq D$. Если переопределить $u_{j_1..j_M}$ как $u_{j_1..j_M}\Delta_{j_1..j_M}$ или на равномерной сетке разбиения положить $\Delta_{j_1..j_M} = 1$, то в правой части уравнения (2) получим выражение дискретной свертки многомерных массивов $u = \{u_{i_1..i_M}\}$ и $v = \{v_{i_1..i_M}\}$.

В одномерном случае когда $x = t$ является временной координатой, ЛИ-уравнение представляет собой интеграл Дюамеля, а ЛИ-система является линейной стационарной, аппаратная функция $u(t)$ которой описывает динамические искажения сигналов. В многомерном случае ЛИ-уравнение представляет собой интегральное уравнение Фредгольма первого рода с ядром типа свертки, а передаточная функция $u(x)$ описывает также пространственные искажения, например, искажения данных при статистической обработке, фотографий при дефокусировке и т.п. Таким образом, ЛИ-системы включают в себя широкий класс систем обработки, инвариантных по отношению к временным и пространственным координатам, и используются в самых разных областях науки и техники, где передаточная функция имеет различные названия – аппаратная функция, функция преобразований или искажений, функция наложений или суперпозиции, ядро преобразований или искажений, и т.д.

2. Метод тестовых испытаний

В общем случае, передаточные функции ЛИ-систем принадлежат к классу обобщенных, что, с учетом, как правило, некорректно поставленных задач реконструкции и нестабильности решений, создает серьезные трудности как при их определении, так и при использовании в методах математической реконструкции сигналов. Для решения этих проблем в работах [8,9] был предложен метод тестовых испытаний, в основе которого лежит уравнение тестовых испытаний (ТИ-уравнение):

$$(v_c * w)(x) - (w_c * v)(x) = \{\eta_T(x)\}, \qquad (3)$$

получающееся из ЛИ-уравнений (1) для обрабатываемых сигналов $v(x)$, $w(x)$ и тестовых сигналов $v_c(x)$, $w_c(x)$ при вычитании их сверток с функциями $v_c(x)$ и $v(x)$ и исключении передаточной функции. При этом неопределенность ТИ-уравнения (3): $\eta_T(x) = \big(v_c * (\Delta w - u_0 * \Delta v) - v * (\Delta w_c - u_0 * \Delta v_c)\big)(x)$ оказывается меньше, чем неопределенности $\eta(x)$ и $\eta_c(x)$ исходных ЛИ-уравнений за счет сокращения членов, входящих в $\eta(x)$ и $\eta_c(x)$ с противоположными знаками. Кроме того, ТИ-уравнение не содержит передаточных функций и при его использовании не возникает ошибок представления функций из класса обобщенных. Все это приводит к снижению ожидаемой погрешности реконструкции сигналов из ТИ-уравнения и повышению устойчивости алгоритма восстановления.

В спектральной форме ТИ-уравнение имеет вид: $W_c(\omega)V(\omega) = V_c(\omega)W(\omega) - \{N_T(\omega)\}$, а в операторной форме: $\mathcal{W}_c v(x) = \mathcal{V}_c w(x) - \{\eta_T(x)\}$, где $\mathcal{W}_c$ и $\mathcal{V}_c$ – операторы свертки с производящими функциями $w_c(x)$ и $v_c(x)$. В дискретном виде ТИ-уравнения: $\mathcal{V}_c w = \mathcal{W}_c v - \{\eta_T\}$ многомерные операторы свертки понимаются в смысле (2). На равномерной сетке разбиения элементы объема в обеих частях этого уравнения сокращаются, и в отличие от уравнения (2), переопределение функций здесь не требуется.

Сигналы ЛИ-систем относятся к функциям основного класса, и их определение не вызывает затруднений, поскольку могут быть использованы как численные, так и экспериментальные методы. При экспериментальном определении тестовых сигналов не требуется ни анализ ЛИ-систем, ни учет паразитных параметров их элементов, так как эта информация содержится в тестовых сигналах. Это упрощает процедуру реконструкции и позволяет восстанавливать сигналы высокоскоростных и прецизионных систем обработки практически любой сложности.

ТИ-уравнение является уравнением типа свертки. При этом оно не содержит передаточных функций из класса обобщенных, а лишь сигналы из класса основных функций, что позволяет эффективно применять существующие методы решения некорректно поставленных и плохо обусловленных задач реконструкции сигналов с высоким уровнем шумов и неопределенностей, такие, например, как методы регуляризации, итерационные методы, вейвлет анализ и т.д.

Согласно теории регуляризации (см., напр., [10]), решением $v_r(x)$ ТИ-уравнения будет функция, минимизирующая функционал $\Phi[v] = \|\mathcal{V}_c w - \mathcal{W}_c v\|^2 + \lambda\|\mathcal{P}v\|^2$ со стабилизатором $\lambda\|\mathcal{P}v\|^2$, где $\mathcal{P}$ – оператор ограничений, накладываемых на реконструируемую функцию $v(x)$, а λ – параметр регуляризации. Не накладывая дополнительные ограничения $\mathcal{P}$ на решения ТИ-уравнений, далее будем полагать: $\mathcal{P}^+\mathcal{P} = \mathcal{I}$, где $\mathcal{I}$ – тождественный оператор, что соответствует поиску решений со “степенью непрерывности” определяемой значением λ. Приравнивая к нулю вариацию функционала $\Phi[v]$ по v, получим:

$$v_r(x) = (\mathcal{W}_c^+\mathcal{W}_c + \lambda\mathcal{I})^{-1}\mathcal{W}_c^+\mathcal{V}_c w(x). \qquad (4)$$

Здесь $\mathcal{W}_c^+$ – сопряженно-транспонированный оператор к оператору $\mathcal{W}_c$ с производящей функцией $w_c(x)$. Согласно теории регуляризации остаток $\big\|\big(v_c * w - w_c * v_r(\lambda)\big)(x)\big\|$ монотонно возрастает с увеличением λ, поэтому оптимальное значение λ можно найти из уравнения: $\big\|\big(v_c * w - w_c * v_r(\lambda)\big)(x)\big\| = \|\eta_T(x)\|$. Отсюда следует и необходимое условие для восстановления сигналов из ТИ-уравнения: $\|\eta_T(x)\| \le \|(v_c * w - w_c * v)(x)\|$. В спектральной форме решение (4) имеет вид: $V_r(\omega) = W_c^+(\omega)V_c(\omega)W(\omega)/(W_c^+(\omega)W_c(\omega) + \lambda)$.

Из (4) получим дискретное регуляризированное решение ТИ-уравнения в операторном:

$$v_r = (\mathcal{W}_c^+\mathcal{W}_c + \lambda\mathcal{I})^{-1}\mathcal{W}_c^+\mathcal{V}_c w \qquad (5)$$

и в спектральном виде: $V^r_{j_1..j_M} = W^{c+}_{j_1..j_M}V^c_{j_1..j_M}W_{j_1..j_M}/\big(W^{c+}_{j_1..j_M}W^c_{j_1..j_M} + \lambda\big)$.

Количество операций необходимых для восстановления массивов с n элементами с помощью операторного уравнения (5), основанного на уравнении дискретной свертки (2), может быть

оценено как: $N(n) \sim O(5n^2 + n + 1)$, что для больших n намного быстрее, чем, например, вычисления свертки одномерных сигналов с помощью матриц Тёплица, требующих $\sim 5n^3$ операций. Необходимое количество операций при использовании спектральной формы уравнения (5) (с прямым и обратным ДПФ) равно: $N(n) \sim O(4n \lg(n) + n + 1)$.

Спектральные методы (5) быстрее операторных, но возможности их применения ограничены, поскольку для определения спектров необходимо знать значения функций во всем пространстве, а сигналы, зачастую, известны лишь на ограниченных областях D, за пределами которых их значения не равны нулю и не периодические.

3. Примеры реконструкции и обсуждение

Результаты численного эксперимента по восстановлению одномерных сигналов сложной формы (суперпозиция трапецеидального и гауссового импульсов) на входе емкостного делителя напряжения из искаженного выходного сигнала $w(t)$ по тестовым сигналам $w_c(t)$ и $v_c(t)$ представлены на рисунке 1.

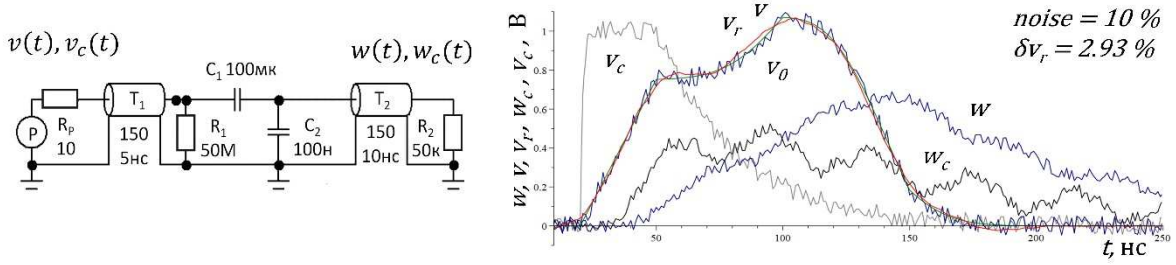


Рисунок 1 – Восстановление динамических сигналов

В качестве тестового воздействия $v_c(t)$ использовался трапецеидальный импульс с разными экспоненциальными фронтами, обеспечивающий широкий спектр сигнала (подробности см. в работах [8,9]). Исходные сигналы были рассчитаны с помощью симулятора аналоговых/цифровых схем Micro-CAP 12.2.0.4 компании Spectrum Software. Принципиальная схема делителя была взята из стандартной библиотеки этой программы. Частота среза делителя составляла 1 МГц, но с помощью математической реконструкции ТИ-методом удалось восстановить наносекундные сигналы с верхней границей спектра больше 1 ГГц. Таким образом, эффективная полоса пропускания системы обработки была улучшена на три порядка. Относительные ошибки реконструкции δv_r при наличии гауссовского шума с дисперсией 5 и 10 % от максимальных значений сигналов v, v_c и w, w_c составили 1.73 и 2.93 %. Численные расчеты проводились в среде Maple 2017.3 Waterloo Maple Inc. на ноутбуке HP 255 G7 с процессором AMD Ryzen 3, 2.5 ГГц. Время восстановления дискретных сигналов размером $n = 501$, соответствующих шагу дискретизации $\Delta = 0.5$ нс составило 1.6..2.0 сек. При размере векторов $n = 251$ ($\Delta = 1.0$ нс) время вычислений составило 0.4…0.5 сек, что соответствует сделанным ранее оценкам квадратичной зависимости времени вычислений от размера данных (см. комментарии к уравнению (5)). Коды программ находятся в открытом доступе в директории автора в библиотеке GitHub по адресу https://github.com/novikov-borodin/data-rec/signal-rec. При использовании программ ссылки на работы автора и директории автора в библиотеке GitHub обязательны.

На рисунке 2 показан результат восстановления $\mathbf{v}_r$ цветного изображения $\mathbf{v}$ размером 500×700 пикселей в формате *jpeg* из изображения $\mathbf{w}$ размера 549×769, размытого с ядром $\mathbf{u}$ размера 50×70, спектральными ТИ-методами. Время восстановления цветного *jpeg* изображения размером 500×700 пикселей, показанного на рисунке 2, составило 56.83 сек, а размером 250×350 пикселей – 15.78 сек. Пересчет времени восстановления $t = t_2 N(n)/N(n_2) = 58.83$ сек показал хорошее соответствие ($\approx 3\%$) с приведенными ранее оценками. Расчеты также проводились в среде Maple 2017.3 Waterloo Maple Inc. Коды программ доступны по адресу https://github.com/novikov-borodin/data-rec/image-rec. При использовании программ ссылки на работы автора и директории автора в библиотеке GitHub обязательны.

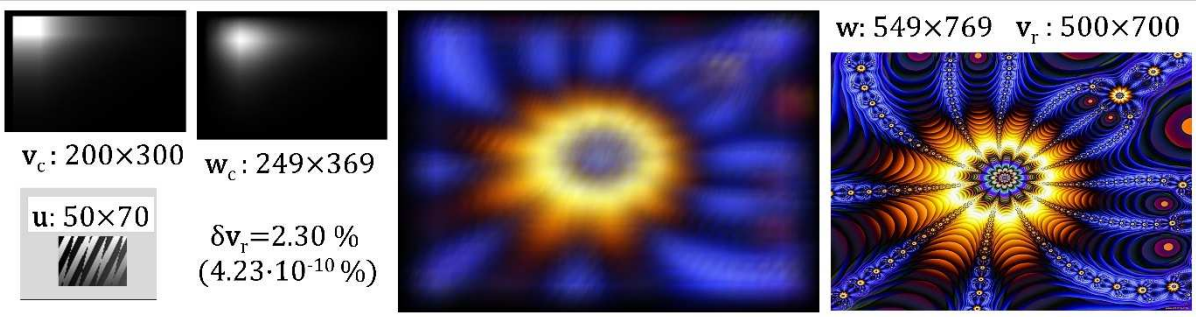


Рисунок 2 – Восстановление изображений

В работе [11] математические методы использовались для расширения полосы пропускания быстродействующего осциллографа путем компенсации его частотной характеристики. Для этого на вход осциллографа подавался сигнал, аппроксимирующий дельта-функцию, и спектр отклика, приближенно совпадающий с импульсной характеристикой осциллографа (его передаточной функцией), использовался для корректировки спектральной характеристики осциллографа и расширения его эффективной полосы пропускания. Такой подход включает как погрешности аппроксимации дельта-функции входным сигналом и импульсной характеристики осциллографа соответствующим выходным сигналом, так и ограничения, связанные с использованием спектрального подхода к восстановлению сигналов.

В работе [12] была предпринята попытка уменьшить погрешность аппроксимации дельта-функции входным сигналом. Для этого на вход осциллографа подавался сигнал в виде ступенчатой функции $\Theta(t)$ с последующим дифференцированием отклика осциллографа для получения его импульсной характеристики. Действительно, согласно свойствам свертки из (1) следует: $w'(t) - \{\eta'(t)\} = (u * \Theta')(t) = (u * \delta)(t) = u(t)$. Такой подход, помимо его сложности, также сопряжен с погрешностями формирования входного сигнала в виде ступеньки и с резким ростом уровня высокочастотных шумов и неопределенности уравнения (1) при дифференцировании, поскольку согласно преобразованию Фурье: $FT[\eta'(t)] = i\omega\, FT[\eta(t)]$. Эксперимент в работе [12] показал, что при увеличении эффективной полосы пропускания осциллографа в 1,5 раза среднеквадратичное значение шума в восстановленном импульсном сигнале увеличивается в 1,5–3 раза. Отметим, что при таком подходе также сохраняются все трудности и ограничения, связанные с использованием спектрального подхода к восстановлению сигналов.

Метод тестовых испытаний позволяет избежать многих проблем, подобных [11,12]. Действительно, в ТИ-методах входной и выходной тестовые сигналы осциллографа могут быть непосредственно использованы для реконструкции сигналов без определения импульсной характеристики или его спектра. При этом, в ТИ-методах не требуется ни спектральных преобразований, приводящих к ограничениям спектрального подхода, ни дополнительных операций, таких как дифференцирование сигналов, приводящих к увеличению уровня шумов и, следовательно, погрешности восстановления.

Заключение

Рассмотрены возможности применения на практике математических методов восстановления многомерных сигналов и изображений, искаженных и зашумленных при их обработке линейными инвариантными системами с помощью тестовых испытаний этих систем. В предлагаемых неслепых методах реконструкции вместо передаточных функций систем обработки, являющихся также функциями искажения сигналов, используются тестовые сигналы систем, что значительно упрощает процедуру восстановления сигналов и делает ее более устойчивой и точной. Операторный подход, основанный на уравнении многомерной свертки, позволяет восстанавливать многомерные сигналы и изображения и существенно повышает скорость численных расчетов. Предложенная методика расширяет возможности применения методов математической реконструкции на практике, позволяет повысить качество обработки сигналов и изображений без сложной и дорогостоящей модернизации оборудования. Методика имеет широкую область практического применения.

Литература

1. Kazufumi I., Bangti J. Inverse Problems: Tikhonov Theory and Algorithms // Singapore: World Scientific, 2014.
2. Chaganova O.B., Grigoryev A.S., Nikolaev D.P., Nikolaev I.P. Applied aspects of modern non-blind image deconvolution methods // Computer Optics, 48(4), 2024. doi: 10.18287/2412-6179-CO-1409.
3. Wang H. et al. Ambulatory ECG noise reduction algorithm for conditional diffusion model based on multi-kernel convolutional transformer // The Review of scientific instruments. 95(9), 2024.
4. Junjie T. A.O., et al. An image defocus deblurring method based on gradient difference of boundary neighborhood // Virtual Reality & Intelligent Hardware, 5(6): 538-549, 2023. doi: 10.1016/j.vrih.2023.06.008.
5. Levin A., et al. Understanding and evaluating blind deconvolution algorithms // 2009 IEEE conference on computer vision and pattern recognition, IEEE, 2009.
6. Hyeontae Jo, Hwijae Son, Hyung Ju Hwang, Eun Heui Kim. (2020). Deep neural network approach to forward-inverse problems. // Networks and Heterogeneous Media, 15(2): 247-259. doi: 10.3934/nhm.2020011.
7. Malay I. Digital Signal Processing in Modern Radio Engineering Measurements // Plenary Talk at XXV International Conference Digital Signal Processing and Its Applications – DSPA-2023, 2023.
8. Novikov-Borodin A.V. Using Test Trials of Processing Systems for Non-Blind Recovery of Signals and Images // Journal of Data Science and Intelligent Systems, 2025. doi: 10.47852/bonviewJDSIS52024990.Новиков-Бородин А.В. Математическая реконструкция сигналов и изображений с помощью тестовых испытаний: неслепой подход. Журнал вычислительной математики и математической физики, т.65, № 3, 2025, с.401-414. doi: 10.31857/S0044466925030135.Elsayed M.A., El-Shafai W., Rashwan M.A., Dessouky M.I., El-Fishawy A.S., Abd El-Samie F.E. Efficient iterative implementation of regularized solutions for image and signal reconstruction problems // Journal of Optics, 2023. doi: 10.1007/s12596-023-01179-w.
11. Zhu J., Wang Y., Zhao K., Wang Y. Oscilloscope Measurement Bandwidth Expansion Using Frequency Response Compensation. IEEE Access, vol. 12: 49772-49778, 2024. doi: 10.1109/ACCESS.2024.3384381.
12. Kleopin A.V., Zenchenko M.A. Tikhonov Regularization in Pulse Signal Processing for Oscilloscope Measurements // XXIV International Conference on Digital Signal Processing and its Applications DSPA-2022, 2022. doi: 10.1109/DSPA53304.2022.9790777.

APPLICATION OF TEST TRIAL METHODS FOR SOLVING THE PRACTICAL TASKS OF RECOVERY OF SIGNALS DISTORTED BY PROCESSING SYSTEMS

A.V. Novikov-Borodin
Institute for Nuclear Research of RAS, 117312 Moscow, Russia

Abstract. This article explores the potential of using the mathematical methods of test trials to solve practical problems of reconstructing one-dimensional and multidimensional signals, including images, distorted during processing by linear invariant systems described by a multidimensional convolution-type equations. Types of invariant systems, including stationary ones, are considered, and examples are provided. The test trial methods enable non-blind signal reconstruction without determining the transfer functions of processing systems belonging to the class of generalized functions. Instead, test signals from the class of ordinary functions are used. The use of test signals significantly simplifies the reconstruction procedure, making it more accurate and robust. An operator approach based on the multidimensional convolution equation increases the speed of numerical computations. To solve ill-posed and ill-conditioned problems, a regularization technique is used, allowing for the efficient reconstruction of non-deterministic signals with high levels of noise and uncertainty. Examples of the methods' application are provided, and the reconstruction procedure is compared with existing approaches. The proposed method allows us to expand the scope of application of mathematical reconstruction methods and improve the quality of signal and image processing without the use of complex and expensive hardware modification methods.
Program codes and datasets are openly available in GitHub library at: https://github.com/novikov-borodin/data-rec.
Keywords: mathematical processing of signals and images; linear invariant systems; test trials of processing systems; multidimensional convolution equation; ill-posed and ill-conditioned problems; deconvolution; regularization technique.

МАТЕМАТИЧЕСКАЯ ОБРАБОТКА МНОГОМЕРНЫХ СИГНАЛОВ И ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ МЕТОДА ПРЯМЫХ СДВИГОВ

с.н.с., к.т.н. Новиков-Бородин А.В.
Институт ядерных исследований РАН, 117312 Москва, Россия

Аннотация: Рассматриваются возможности применения математических методов прямых сдвигов для коррекции, реконструкции и моделирования сигналов, искаженных линейными инвариантными системами в процессе обработки. Линейные инвариантные системы охватывают широкий класс систем обработки, описываемых многомерными уравнениями типа свертки, и являются расширением понятия линейных стационарных систем на пространственные координаты. Методы прямых сдвигов позволяют в ряде важных и часто встречающихся на практике случаях решать задачи обработки сигналов и изображений при минимальной исходной информации с минимальными вычислительными ресурсами и погрешностями. Эффективность метода прямых сдвигов обусловлена возможностью выбора вида преобразований и сокращения их числа. Рассмотрены примеры использования метода прямых сдвигов на практике. Приведены результаты сравнения метода прямых сдвигов с методами Гаусса-Жордана, прямой матричной инверсии, LU- и SVD-разложений.

Программные коды доступны в библиотеке GitHub по адресу: https://github.com/novikov-borodin/lst-data-proc.

Ключевые слова: математическая обработка сигналов и изображений; метод прямых сдвигов; метод тестовых испытаний; линейные инвариантные системы; многомерное уравнение типа свертки.

Введение

Несовершенство систем обработки приводит к появлению искажений в получаемых, преобразуемых и передаваемых ими данных. Аппаратные методы уменьшения искажений связаны с решением сложных научно-технических и технологических задач, а также с созданием дорогостоящего оборудования. Альтернативой им являются математические методы реконструкции неискаженных данных из искаженных в процессе обработки (см., например, [1,2] и ссылки в них). Предлагаемые в данной работе математические методы прямых сдвигов (ПС) предназначены для коррекции, реконструкции и моделирования сигналов и изображений, искаженных в процессе их обработки линейными инвариантными (ЛИ) системами. ЛИ-системы охватывают широкий класс систем обработки, описываемых многомерными уравнениями типа свертки, и являются расширением понятия линейных стационарных систем (см., например, [3]) на пространственные координаты.

Методы прямых сдвигов позволяют решать задачи обработки сигналов при минимуме исходной информации с минимальными погрешностями и вычислительными ресурсами. Эффективность ПС-метода обусловлена алгоритмом его реализации с возможностью выбора вида преобразований и сокращения их числа. Метод прямых сдвигов был разработан автором для решения частных задач коррекции данных нейтронного спектрометра по времени пролета ИЯИ РАН [4] и для моделирования зон облучения новообразований на комплексе протонной терапии ИЯИ РАН [5]. В данной работе проанализированы достоинства и недостатки ПС-методов, рассмотрены возможности их применения для решения целого ряда часто встречающихся на практике типовых задач. Приводятся результаты сравнения ПС-метода с другими методами прямой инверсии, такими как метод Гаусса-Жордана, прямой матричной инверсии, LU- и SVD- разложений.

1. Метод прямых сдвигов

Функциональную систему обработки одномерных или многомерных сигналов, включая изображения, будем называть линейной инвариантной, если в области $x = (x_1, x_2, \dots) \in D \subset \mathbb{R}^n$ она характеризуется передаточной функцией $u(x)$ и ее реакция $w(x)$ (выходной сигнал) на внешнее воздействие $v(x)$ (входной сигнал) описывается многомерным уравнением типа свертки, которое будем называть уравнением ЛИ-систем, или, кратко, ЛИ-уравнением:

$$w(x) - \{\eta(x)\} = (u * v)(x) = \int_D u(x-\xi)v(\xi)\mathrm{d}\xi. \qquad (1)$$

Неопределенность ЛИ-уравнения: $\eta(x) = (\Delta w - u * \Delta v - \Delta u * v + \Delta u * \Delta v)(x)$ можно определить из уравнения свертки: $w_0(x) = (u_0 * v_0)(x)$ детерминированных функций без неопределенностей: $w_0(x) = (w - \Delta w)(x)$, $u_0(x) = (u - \Delta u)(x)$ и $v_0(x) = (v - \Delta v)(x)$.

Вводя оператор свертки $\mathcal{F}$ с порождающей функцией $f(x)$: $\mathcal{F}g(x) = (f * g)(x)$, ЛИ-уравнение можно представить в операторной форме: $\mathcal{U}v(x) = w(x) - \{\eta(x)\}$, а используя теорему Бореля о свертке, – в спектральной форме: $U(\omega)V(\omega) = W(\omega) - \{N(\omega)\}$, где $F(\omega) = FT[\mathcal{F}] = FT[f(x)]$ – спектр (преобразование Фурье) функции $f(x)$ и оператора $\mathcal{F}$, и $N(\omega) = \Delta W(\omega) - U(\omega)\Delta V(\omega) - \Delta U(\omega)V(\omega) - \Delta U(\omega)\Delta V(\omega)$. Спектральная форма ЛИ-уравнения имеет простой вид, но для определения спектров функции должны быть известны во всем пространстве $\mathbb{R}^n$, а не только в заданной области D, что резко ограничивает область применения спектральных методов.

ЛИ-уравнение в дискретном виде представляет собой частичную сумму интеграла (1):

$$w_{i_1..i_M} - \{\eta_{i_1..i_M}\} = \sum_{j_1=1}^{m_1} \cdots \sum_{j_M=1}^{m_M} u_{j_1..j_M}\, v_{i_1-j_1..i_M-j_M} \Delta_{j_1..j_M}, \qquad (2)$$

где $w = \{w_{i_1..i_M}\}$, $\eta = \{\eta_{i_1..i_M}\}$, $u = \{u_{i_1..i_M}\}$ и $v = \{v_{i_1..i_M}\}$ – M-мерные массивы, соответствующие функциям $w(x)$, $\eta(x)$, $u(x)$ и $v(x)$, $\Delta_{j_1..j_M}$ – элемент объема, представляет собой частичную сумма интеграла (1) на сетке $\{d_{i_1....i_M}\}$, $i_k = 1 \dots m_k, k = 1 \dots M$, разбивающей прямоугольную область $D_0 \supseteq D$. Если переопределить $u_{j_1..j_M}$ как $u_{j_1..j_M}\Delta_{j_1..j_M}$ или на равномерной сетке разбиения положить $\Delta_{j_1..j_M} = 1$, то в правой части уравнения (2) получим выражение дискретной свертки многомерных массивов $u = \{u_{i_1..i_M}\}$ и $v = \{v_{i_1..i_M}\}$.

Сдвиг функции $f(x)$ можно осуществить с помощью оператора свертки $\mathcal{S}$ с порождающей функцией $s(x)$: $f(x - x_0) = \mathcal{S}f(x) = (s * f)(x)$, где $s(x) = \delta(x - x_0)$ и $\delta(x)$ – дельта-функция Дирака. Согласно линейности, коммутативности и ассоциативности свертки последовательное применение $\mathcal{S} = \mathcal{S}_n \dots \mathcal{S}_2\mathcal{S}_1$ и линейные комбинации $\mathcal{S} = \sum_i a_i\mathcal{S}_i$ операторов сдвига являются операторами свертки. Осуществляя преобразования $\mathcal{S}$ над обеими частями ЛИ-уравнения, получим уравнение прямых сдвигов:

$$\mathcal{S}w(x) - \{\mathcal{S}\eta(x)\} = (s * u * v)(x) = (\mathcal{S}u * v)(x) = (u * \mathcal{S}v)(x). \qquad (3)$$

Согласно (3), осуществляя ПС-преобразования над обеими частями ЛИ-уравнения можно из отклика $w(x)$ ЛИ-системы на входной сигнал $v(x)$ смоделировать ее отклик $\mathcal{S}w(x)$ на входной сигнал $\mathcal{S}v(x)$, не определяя ее передаточную функцию; или же отклик $\mathcal{S}w(x)$ системы с передаточной функцией $\mathcal{S}u(x)$ на такой же входной сигнал, не определяя его. При этом ошибка преобразований составит $\{\mathcal{S}\eta(x)\}$.

2. Базовая задача ПС-преобразований

В основе ПС-методов лежит базовая задача о наложении двух импульсов, к которой сводится множество типовых задач, часто встречающихся на практике. Примеры подобных задач будут приведены ниже. Рассмотрим решение этой задачи для одномерных сигналов ЛИ-систем. В многомерном случае решение аналогично.

Задача (наложение двух импульсов). Пусть известен отклик $w(x)$ ЛИ-системы (1) от двух смещенных относительно друг друга на расстояние (или время) $T > 0$ импульсов $v_0(x)$: $v(x) = v_0(x) - av_0(x - T)$ ($w(x) = v(x) = 0$ при $x < 0$). Реконструировать на интервале $x \in (-\infty, L)$, $L > T$, отклик $w_0(x)$ этой системы от импульса $v_0(x)$.

Из ЛИ-уравнения (1) имеем: $w(x) = (u * s * v_0)(x) = (s * w_0)(x)$, $s(x) = \delta(x) - a\delta(x - T)$. Применяя к функции $s(x)$ последовательность ПС-преобразований: $\mathcal{S} = \mathcal{S}_n \cdots \mathcal{S}_1$, где $\mathcal{S}_k$: $\delta(x) + a^{2^{k-1}}w(x - 2^{k-1}T)$, получим: $\mathcal{S}s(x) = \delta(x) - a^{2^n}\delta(x - 2^nT)$. Тогда, согласно (2): $\mathcal{S}w(x) = (\mathcal{S}s * w_0)(x) = w_0(x) - a^{2^n}w_0(x - 2^nT)$. При $2^nT > L$ на интервале $x \in (-\infty, L)$ имеем: $\mathcal{S}w(x) = w_0(x)$ (см. рисунок 1) и задача решена.

При наличии неопределенностей $\eta(x)$ погрешность $\mathcal{S}_k \cdots \mathcal{S}_1\eta(x)$ на интервале $[2^{k-1}T, 2^kT)$ не превысит: $\eta(x)\left(1 + \sum_{i=1}^{k} a^{2^i}\right)$. При хорошей обусловленности задачи ($|a| < 1$) прирост погрешности a^{2^k} быстро уменьшается, а при плохой ($|a| > 1$) – возрастает (см. Таблицу 1).

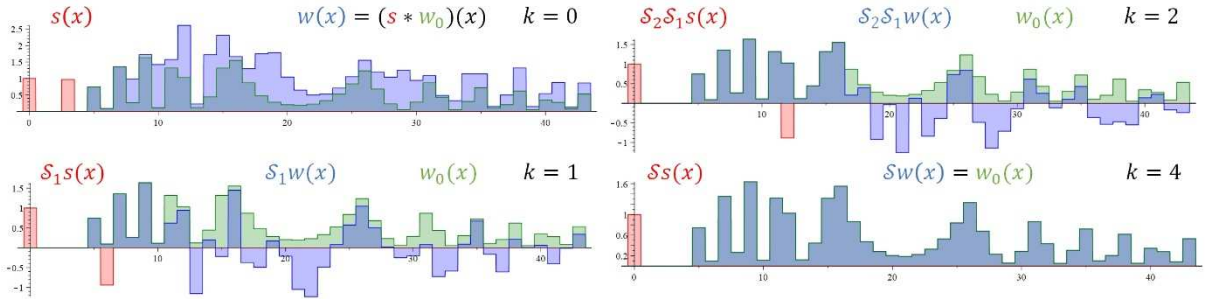


Рисунок 1 – Реконструкция сигналов при наложении двух импульсов

Таблица 1 – Прирост погрешности $\boldsymbol{a^{2^k}}$ на интервале $[\boldsymbol{2^{k-1}T}, \boldsymbol{2^kT})$ на $\boldsymbol{k}$-том шаге итерации

$k \setminus a$	**0.5**	**0.7**	**0.9**	**1.1**	**1.3**	**1.5**
1	0.25	0.49	0.81	1.21	1.69	2.25
3	$3.9\ 10^{-3}$	$5.8\ 10^{-2}$	$4.3\ 10^{-1}$	2.14	8.16	25.6
5	$2.3\ 10^{-10}$	$1.1\ 10^{-5}$	$3.4\ 10^{-2}$	21.1	$4.4\ 10^{3}$	$4.3\ 10^{5}$
7	$2.9\ 10^{-39}$	$1.5\ 10^{-20}$	$1.4\ 10^{-6}$	$1.9\ 10^{5}$	$3.8\ 10^{14}$	$3.5\ 10^{22}$
9	$7.5\ 10^{-155}$	$4.9\ 10^{-80}$	$3.7\ 10^{-24}$	$1.6\ 10^{21}$	$2.2\ 10^{58}$	$1.4\ 10^{90}$

Таким образом за $n = [log_2(L/T)] + 1$ итераций (или меньше, когда приростом погрешности можно пренебречь) можно на произвольном интервале $x \in (-\infty, L)$ реконструировать отклик любой ЛИ-системы от одного импульса при наложении двух смещенных входных импульсов. При $v_0(x) = \delta(x)$ будет реконструирована передаточная функция $u(x)$ ЛИ-системы. #

В случае многомерных сигналов функции рассматриваются в области $x = \{x_i\} \in D \subset \mathbb{R}^n$, а смещение задается вектором $\boldsymbol{T} = \{T_i\} \in \mathbb{R}^n$. На рисунке 2 представлен пример решения базовой задачи при наложении двумерного изображения в случае плоскопараллельного смещения.

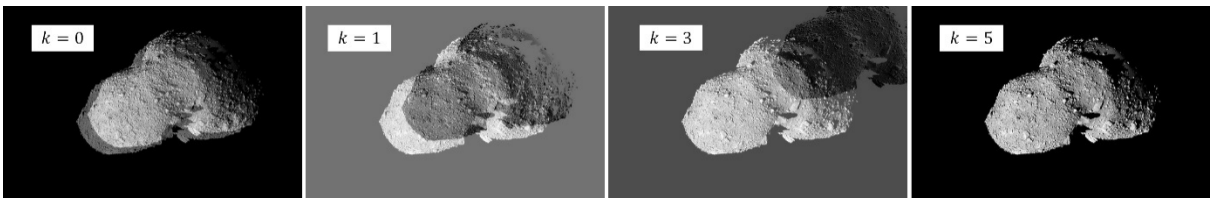


Рисунок 2 – Реконструкция при наложении двух изображений

В многомерном случае не все смещения являются ПС-преобразованиями. Например, комбинации вращений и плоскопараллельных смещений не являются ПС-преобразованиями.

3. Алгоритм численных вычислений

Численные вычисления с помощью ПС-преобразований можно реализовать на основе решения задачи о произвольных дискретных наложениях сигналов.

Задача (дискретные наложения сигналов). Из отклика $w(x)$ ЛС-системы на входной сигнал $v(x) = (s * v_0)(x)$, где $s(x) = \sum_{i=0} s_i \delta(x - it)$, $s_0 \neq 0$, $t > 0$, $w(x) = v(x) = 0$ при $x < 0$, реконструировать на интервале $x \in (-\infty, L)$ отклик $w_0(x)$ от импульса $v_0(x)$.

Применяя к функции $s(x)$ ПС-преобразования $\mathcal{S} = \mathcal{S}_n \cdots \mathcal{S}_1$, где $\mathcal{S}_k$: $\delta(x) - (s_k^{k-1}/s_0)\delta(x - kt)$ и $\mathcal{S}_k \cdots \mathcal{S}_1 s(x) = s^k(x) = \sum_{i=0} s_i^k \delta(x - it)$, $s^0(x) = s(x)$, последовательно обнуляем

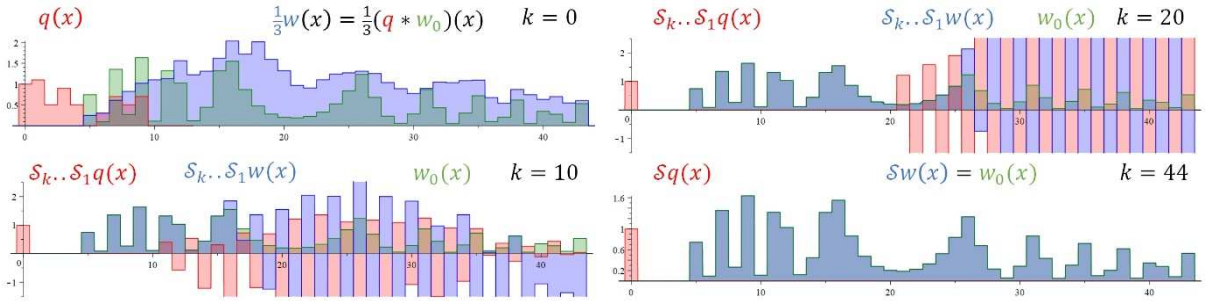


Рисунок 3 – Реконструкция сигналов при дискретных наложениях импульсов

коэффициенты s_k^k в функциях $s^k(x)$ (см. рисунок 2). На n-том шаге итерации при $nt > L$ на интервале $x \in (-\infty, L)$ получим: $s^n(x) = \mathcal{S}s(x) = s_0\delta(x)$. Отсюда, согласно (2): $\mathcal{S}w(x) = s_0 w_0(x)$. Погрешность решения на интервале $x \in [kt, kt + t)$ составит: $\mathcal{S}_k \cdots \mathcal{S}_1\eta(x)/s_0$. #

Преимущество данного алгоритма заключается в том, что если на k-том шаге итерации $\left|s_k^{k-1}\right| < \varepsilon$, где ε – ошибка вычислений при численных расчетах или при наличии шумов и неопределенностей – ожидаемый уровень шумов в данных, то этот шаг можно пропустить и перейти к следующему, что значительно сокращает время вычислений, и снижает погрешность, т.к. неопределенности на пропущенных шагах итерации не добавляются.

В таблице 2 представлено время вычислений в секундах и отношения времен реконструкции сигналов длительности n при решении базовой задачи с помощью описанного алгоритма метода прямых сдвигов (ПС); обратного хода метода Гаусса-Жордана (ГЖ) $v_{i+1} = \left(w_i - \sum_{j=1}^{i} s_{i-j}w_j\right)/s_0$, $i = 0..n$; прямой матричной инверсии (ПМ) $v = \boldsymbol{S}^{-1}w$; а также LU- и SVD- методами, использующими соответствующие разложения матрицы $\boldsymbol{S}$. Расчеты проводились в среде Maple 2017.3 Waterloo Maple Inc. на ноутбуке HP 255 G7 с процессором AMD Ryzen 3, 2.5 ГГц. Использовались ПМ, LU и SVD методы из стандартных пакетов *LinearSolve* и *LinearAlgebra* программы Maple. Время вычислений с помощью метода прямых сдвигов значительно меньше времени вычислений другими методами. Например, при длине векторов $n = 6400$ время расчета с помощью ПС-метода в 2707 раз меньше времени расчета SVD-методом. Кроме того, при проведении расчетов ПС-методом никаких дополнительных массивов или матриц кроме исходных данных не требуется, что значительно экономит вычислительные ресурсы.

Таблица 2 – Время решения базовой задачи (сек) при различной длине сигналов $\boldsymbol{n}$

n	ПС	ГЖ	ПМ	LU	SVD
200	0.015	0.032 (2.13 ПС)	0.140 (9.33 ПС)	0.297 (19.8 ПС)	0.250 (16.7 ПС)
400	0.047	0.094 (2.00 ПС)	0.578 (12.3 ПС)	0.891 (19.0 ПС)	1.125 (23.9 ПС)
800	0.062	0.266 (4.29 ПС)	2.375 (38.3 ПС)	2.750 (44.4 ПС)	4.375 (70.6 ПС)
1600	0.094	1.250 (13.3 ПС)	11.38 (121.0 ПС)	11.70 (124.5 ПС)	22.47 (239.0 ПС)
3200	0.266	6.109 (23.0 ПС)	53.656 (201.7 ПС)	55.234 (207.6 ПС)	133.4 (501.5 ПС)
6400	0.719	23.97 (33.3 ПС)	294.06 (409.0 ПС)	329.44 (458.2 ПС)	1946.5 (2707.2 ПС)

Программные коды численных расчетов ПС-методом находятся в открытом доступе в библиотеке GitHub по адресу: https://github.com/novikov-borodin/lst-data-proc. При использовании программ ссылки на работы автора и директории автора в библиотеке GitHub обязательны.

4. Типовые задачи и примеры использования ПС-методов на практике

Преимущества ПС-методов по времени вычислений при решении базовой задачи проявляются при решении типовых задач, которые можно свести к базовой задаче, что проиллюстрировано на рисунке 4 при реконструкции изображений с равномерным (а) и произвольным (б) размытием. При равномерном размытии уже после первой операции сдвига задача переходит в базовую, и реконструировать неискаженное изображение удается всего за четыре итерации. В случае размытия произвольного вида задача к базовой не сводится и

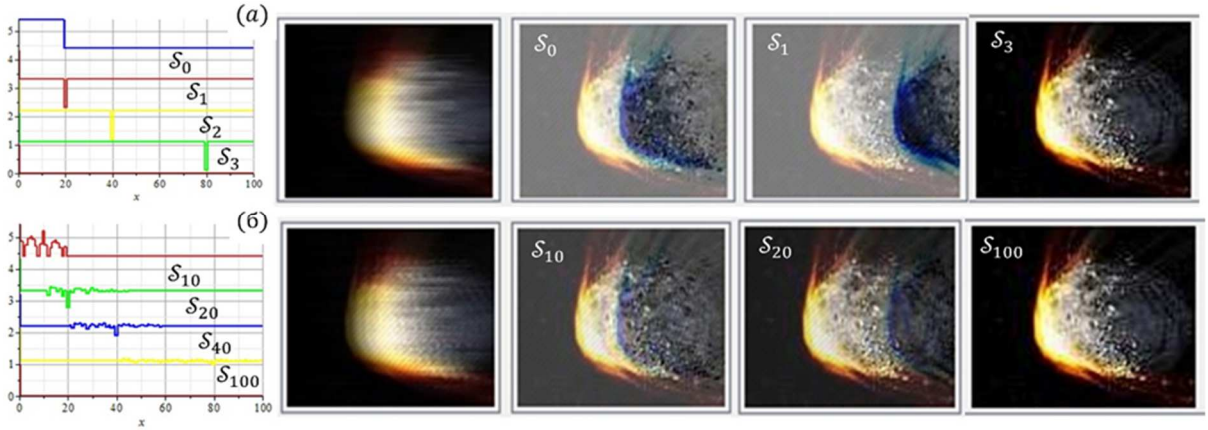


Рисунок 4 – Реконструкция изображений при равномерном (а) и произвольном (б) размытии

приходится проходить все изображение, для чего потребовалось 100 итераций. Тем не менее даже в этом случае никакие дополнительные массивы ПС-методам не требуются, и вычисления осуществляются с помощью сдвигов исходного изображения.

Аналогично можно осуществить реконструкцию отклика ЛИ-системы от прямоугольного импульса или трапецеидального импульса с равными фронтами из импульса большей длительности, т.к. сдвигая сигналы и вычитая их из исходных, задача сводится к базовой.

Решение базовой задачи непосредственно использовалось в работе [4] для коррекции данных нейтронного спектрометра по времени пролета в Институте ядерных исследований РАН, где на основной импульс накладывался послеимпульс, что приводило к искажению спектров. Используя аналогичные рисунку 4 ПС-преобразования можно также смоделировать спектр подобных спектрометров от импульса меньшей длительности, что означает повышение их разрешающей способности без изменения конструкции. В работе [5] ПС-преобразования использовались для моделирования заданных зон облучения $w(x)$ новообразований при протонной терапии. Под $v(x)$ понимался пик Брэгга, а реконструируемая функция $u(x)$ представляла собой искомый план облучения.

На практике передаточные функции систем обработки часто описываются экспоненциальными функциями вида: $u(x) = Aexp(-x/B)$, $x > 0$. В дискретном виде с шагом разбиения Δ их можно представить как $u = A\sum_{i=0} \delta_i exp(-i\Delta/B)$, где δ_i – символ Кронекера. Всего за одну итерацию $\mathcal{S}_0$: $\delta_0 - \delta_1 exp(-\Delta/B)$ функцию u можно преобразовать в единичный импульс $\mathcal{S}_0 u = A\delta_0$ и реконструировать сигнал на входе из искаженного сигнала на выходе: $\mathcal{S}_0 w = (\mathcal{S}_0 u * v) = Av$ (см. рисунок 5а). Например, в работах [6,7] передаточная функция ФЭУ представлялась в виде двух экспонент $u = A\sum_{i=0} \delta_i [exp(-i\Delta/B) - exp(-i\Delta/b)]$, $B > b > 0$, что позволило за два ПС-преобразования $\mathcal{S}_0$: $\delta_0 - exp(-\Delta/B)\delta_1$ и $\mathcal{S}_1$: $C[\delta_0 - exp(-\Delta/B)\delta_1]$, $C = [exp(-\Delta/B) - exp(-\Delta/b)]^{-1}$, реконструировать сигнал на входе из искаженного сигнала на выходе: $\mathcal{S}_1\mathcal{S}_0 w = (\mathcal{S}_1\mathcal{S}_0 u * v) = Av$ (см. рисунок 5б). Это решение используется как для регистрации, так и для разделения сигналов в детекторах частиц большой интенсивности.

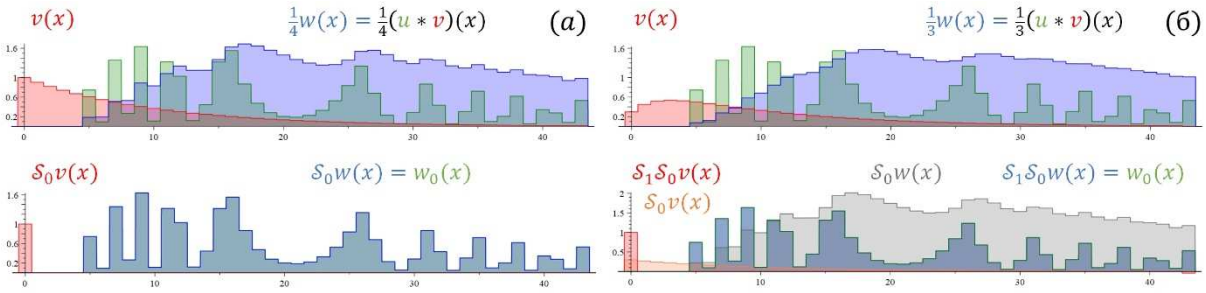


Рисунок 5. Обработка экспоненциальных импульсов

Решения рассмотренных типовых задач для прямоугольных и экспоненциальных импульсов можно использовать для достижения сверх разрешения сигналов и изображений. Последовательно группируя n сигналов с разрешением T, смещенных относительно друг друга на шаг $t = T/n$, можно составить ЛИ-уравнение (1), решением которого и будет сигнал с разрешением t. Причем для экспоненциальных импульсов сверх разрешение достигается всего за одну итерацию.

Типовые задачи плохо обусловлены, если при сведении их к базовой: $|s_0| < |s_1|$ (см. Таблицу 1), что приводит к расходимости решения. Например, в случае трапецеидального сигнала с экспоненциальными фронтами на Рисунке 5а, относительная погрешность реконструкции импульсного отклика $\|u_r - u\|/\|u\|$ с помощью описанного алгоритма составляет 5688 %. Однако, если применить вспомогательное ПС-преобразование дифференцирования $\mathcal{S}$: d/dx к уравнению $w(x) = (u * v)(x)$ и решать задачу: $\mathcal{S}w(x) = (u * \mathcal{S}v)(x)$, то из-за улучшения обусловленности задачи погрешность ее решения не превысит $1.52 \cdot 10^{-12}$ % (см. рисунок 5б).

Рассмотренные типовые задачи и применения ПС-преобразований не являются единственными. Существует множество других преобразований и задач, решения которых можно свести к решению базовой задачи. К недостаткам ПС-методов следует отнести типичные

недостатки методов прямой инверсии – чувствительность к плохой обусловленности задач, когда необходимо применять приближенные методы решения.

Заключение

Рассматриваемые в данной работе математические методы прямых сдвигов предназначены для коррекции, реконструкции и моделирования сигналов и изображений, искаженных в процессе обработки линейными инвариантными системами, удовлетворяющими многомерному уравнению типа свертки. ПС-методы практически не ограничивают класс обрабатываемых функций и в целом ряде важных и часто встречающихся на практике случаях позволяют решать задачи обработки сигналов при минимуме исходной информации с минимальными погрешностями и вычислительными ресурсами, когда многие из существующих как прямых, так и приближенных математических методов оказываются неэффективными. Эффективность ПС-метода обусловлена алгоритмом его реализации с возможностью выбора вида преобразований и сокращения их числа. ПС-методы имеют широкие перспективы практического применения и являются доступной альтернативой аппаратным методам, связанным с решением сложных научно-технических проблем и с созданием дорогостоящего оборудования.

Литература

1. Mueller J.L., Siltanen S. Linear and Nonlinear Inverse Problems with Practical Applications // SIAM, vol.10, 2012.
2. Hansen P.C. Discrete Inverse Problems: Insight and Algorithms. Fundamentals of Algorithms // Philadelphia: SIAM, 2010.
3. Hespanha J.P. Linear System Theory // Princeton university press, 2009.
4. Новиков-Бородин А.В. Реконструкция экспериментальных данных спектрометров по времени пролёта методом пошагового сдвига // Приборы и техника эксперимента, 2018, № 6, с. 17–24. https://doi.org/10.1134/S0032816218050269.
5. Novikov-Borodin A.V. Direct Reconstruction of the Experimental Data in the Case of Ill-Conditioned Problems and in the Presence of Data Distortions // Instruments and Experimental Techniques, Vol.63, No.5, pp.658-669, © Pleiades Publishing, Ltd., 2020. doi: https://doi.org/10.1134/S0020441220050322.
6. Jordanov V.T. Unfolding-synthesis technique for digital pulse processing, part 1: Unfolding // Nucl. Instrum. Methods Phys. Res. A 805, 63-71, 2016. doi: https://doi.org/10.1016/j.nima.2015.07.040
7. Jordanov V.T., Jordanova K.V. Unfolding-synthesis technique for digital pulse processing, Part 2: Synthesis // Nucl. Instrum. Methods Phys. Res. A 1004, 167421, 2022. https://doi.org/10.1016/j.nima.2022.167421.

APPLICATION OF DIRECT SHIFT METHODS FOR SOLVING THE PRACTICAL PROBLEMS OF SIGNAL PROCESSING

A.V. Novikov-Borodin
Institute for Nuclear Research of RAS, 117312 Moscow, Russia

Abstract. The possibilities of using mathematical methods of direct shifts for correction, reconstruction and modeling of signals distorted during processing by linear invariant systems are considered. Linear invariant systems cover a wide class of processing systems described by multidimensional convolution-type equations and are an extension of the concept of linear stationary systems to spatial coordinates. Direct shift methods make it possible, in a number of important and frequently encountered practical cases, to solve signal and image processing problems with minimal initial information with minimal computational resources and errors. The effectiveness of the direct shift method is due to the possibility of choosing the type of transformations and reducing their number. Examples of using the direct shift method in practice are considered. The results of comparing the method of direct shifts with the methods of Gauss-Jordan, direct matrix inversion, and LU and SVD decompositions are presented.

Program codes are openly available in GitHub library at: https://github.com/novikov-borodin/lst-data-proc.

Keywords: mathematical signal and image processing; direct shift techniques; linear invariant systems; multidimensional convolution type equation.

ПОИСК ОБЪЕКТА ИНТЕРЕСА НА МАЛОКОНТРАСТНЫХ ИЗОБРАЖЕНИЯХ С ПОМОЩЬЮ ПОРОГОВОЙ СЕГМЕНТАЦИИ НА ОСНОВЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА

студ. Буянкина Я.А., к.т.н. доц. Бабаян П.В.

Рязанский государственный радиотехнический университет имени В.Ф. Уткина

Актуальная задача сегментации изображений, особенно в случаях, когда гистограмма объекта обладает несколькими пиками (мультимодальность), является предметом данного исследования. Для эффективного решения этой проблемы был выбран и реализован генетический алгоритм. В рамках работы представлена детальная реализация алгоритма, а также проведены экспериментальные исследования с использованием различных изображений. Анализ полученных результатов, включая расчет ключевых характеристик, подтверждает превосходство предложенного генетического подхода над традиционными методами сегментации, такими как метод Отцу, а также энтропийными методами Капура и Тсаллиса.

I. Введение

В области обработки изображений сегментация выделяет интересующие объекты из цифровых кадров. Простой подход — пороговая сегментация: гистограмма яркости делится на части с глобальным порогом, после чего пиксели классифицируются как объект или фон. Эффективность зависит от разделимости гистограммы.

Проблема пороговой сегментации в современной обработке изображений заключается в том, что этот метод может быть недостаточно эффективным при работе с изображениями, имеющими сложные или неоднородные гистограммы яркости.

Для улучшения результатов сегментации в таких случаях часто требуется применение более сложных и вычислительно затратных методов, таких как адаптивная пороговая обработка или использование алгоритмов машинного обучения.

Работа фокусируется на сегментации изображений с тёмным фоном и контрастным объектом (несколько зон яркости), таких как ночные снимки летательных аппаратов, 3D-сканы или ИК-изображения с бликами. При наблюдении таких объектов в неравномерном или малом освещении их изображения содержат зоны разных уровней яркости, зеркальные или блестящие поверхности, блики. Применение традиционных методов сегментации к таким изображениям обычно не обеспечивает приемлемого качества. В настоящей работе для сегментации изображений, относящихся к рассматриваемому классу, предлагается использовать генетический алгоритм [1].

II. Классические методы и их недостатки

С помощью классического метода Отцу [2] вычисляется порог t, минимизирующий среднюю ошибку сегментации, т.е. среднюю ошибку от принятия решения о принадлежности пикселей изображения объекту или фону. Значения яркостей пикселей можно рассматривать как случайные величины, а их гистограмму – как оценку плотности распределения вероятностей. Метод Отцу ищет порог, уменьшающий внутриклассовую дисперсию.

Метод Капура относится к энтропийным методам бинаризации изображений [3]. Энтропия Капура — это метрика для многоуровневой сегментации изображений, которая разделяет изображение на разные классы, а величина энтропии определяет, является ли класс однородным.

Недавние разработки статистической механики, основанные на концепции неэкстенсивной энтропии, также называемой энтропией Тсаллиса, усилили интерес к исследованию возможного распространения энтропии Шеннона на алгоритмы обработки изображений. Энтропия Тсаллиса основана на сегментации изображений с помощью принципа максимальной энтропии [4].

Каждый из этих методов имеет свои особенности и применяется в зависимости от задачи. Однако все они обладают определёнными недостатками и ограничениями, которые необходимо учитывать при выборе метода для конкретной задачи обработки изображений [5].

1. Чувствительность к равномерности распределения интенсивности пикселей и к шуму.

2. Возможность неправильного выбора порога при наличии нескольких пиков в гистограмме.
3. Методы могут быть неэффективны для изображений с низкой контрастностью, сложными текстурами, неоднородным фоном, перекрывающимися объектами и т.д.

Исходя из вышеописанных пунктов можно сделать вывод, что данные методы пороговой сегментации демонстрируют недостаточную эффективность при обработке изображений с тёмным фоном и малоконтрастными объектами интереса [6]. Для решения подобных задач требуется применение альтернативных методов сегментации, более точно учитывающих специфику изображений с указанными характеристиками. В данной работе был применён генетический алгоритм в силу его особенностей:

1. Гибкость (настройка параметров в зависимости от условий).
2. Способность находить глобальные оптимумы.
3. Устойчивость к шуму в изображениях.
4. Возможность многоклассовой сегментации.

III. Генетический алгоритм

Генетический алгоритм (GA) - это естественная метаэвристика, которая имитирует гены. Генетический алгоритм используется для поиска оптимальных решений, когда не существует детерминированного метода или если детерминированный метод сложен в вычислительном отношении. Количество решений на итерации называется размером популяции. Каждое решение представлено в виде хромосомы, и каждая хромосома состоит из генов [7]. Блок-схема алгоритма представлена на рис.1.

Генетический алгоритм начинается с инициализации популяции, состоящей из N случайных решений, каждое из которых представляет собой набор хромосом. Хромосомы, в свою очередь, состоят из генов, кодирующих параметры решаемой задачи. После инициализации выполняется оценка приспособленности каждого решения, что позволяет определить их качество. На основе этих оценок отбираются лучшие решения для дальнейшего использования.

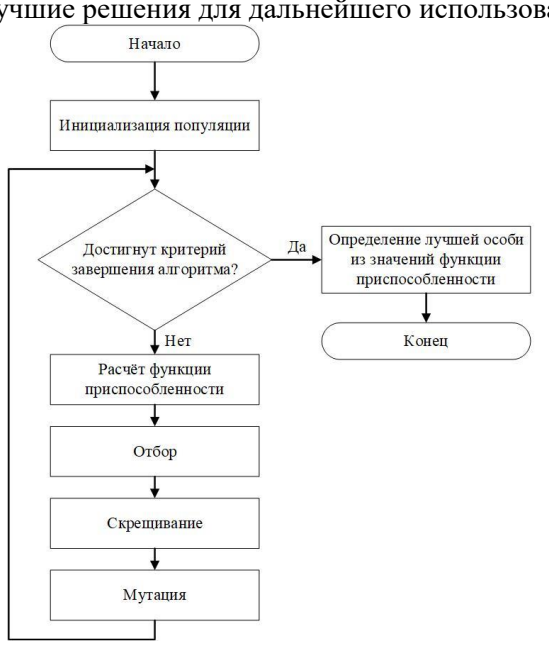


Рисунок 1. Блок-схема общего генетического алгоритма

Процесс воспроизведения состоит из четырех этапов: отбор, скрещивание, мутация и принятие решения.

1. Отбор. Выбираются лучшие кандидаты из популяции, с возможностью включения менее приспособленных для сохранения разнообразия. Популярный метод – "колесо рулетки", где вероятность выбора пропорциональна приспособленности.

2. Скрещивание путем выбора случайной точки в хромосомах и обмена генами после этой точки.
3. Мутация. Для преодоления локальных оптимумов используется мутация, случайным образом изменяющая ген [8].
4. Принятие решения. Новые хромосомы включаются в популяцию, если они лучше родителей.

Завершающий этап генетических алгоритмов заключается в остановке итерационного процесса при достижении определенного критерия. Наиболее распространенным критерием остановки является количество итераций.

Основная задача многопороговой сегментации найти оптимальные пороги, которые максимально точно разделят изображение на нужные сегменты. Поиск таких порогов — это задача оптимизации, которая может быть сложной из-за большого пространства возможных решений, особенно при большом количестве порогов. Для этого и применяется генетический алгоритм, который рассматривает набор порогов как индивидуум (хромосому) и с помощью операций селекции, скрещивания и мутации ищет наиболее подходящие комбинации порогов, а функция приспособленности является критерием, по которому оценивается качество каждого набора порогов.

IV. Реализация многопороговой сегментации

В качестве вычислительной среды выполнения использовался пакет прикладных программ Matlab версии R2021b. На вход программы подаётся исходное малоконтрастное изображение (рис. 2). Рядом приведено осветлённое с помощью гаммы-коррекции изображение для понимания сути исследования (рис. 3). Также подаются параметры, задаваемые пользователем: количество порогов, итераций, размер популяции.



Рисунок 2. Оригинальное изображение



Рисунок 3. Осветлённое изображение

Далее алгоритм вычисляет функции приспособленности, проводит отбор, скрещивание, мутацию. По достижении количества итераций значения функции приспособленности максимизируют, и изображение сегментируется (рис. 4). Также на экран выводятся отдельные области пороговой сегментации (Рис. 5-8, где T – порог).

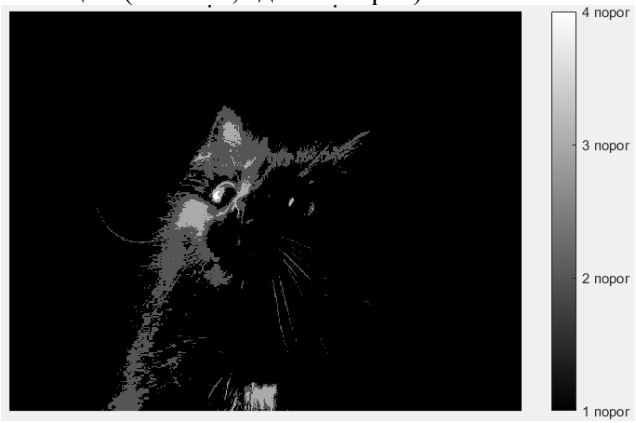


Рисунок 4. Сегментированное изображение



Рисунок 5. Область пороговой сегментации №1 (от 0 до T1)



Рисунок 6. Область пороговой сегментации №2 (от T1 до T2)



Рисунок 7. Область пороговой сегментации №3 (от T2 до T3)

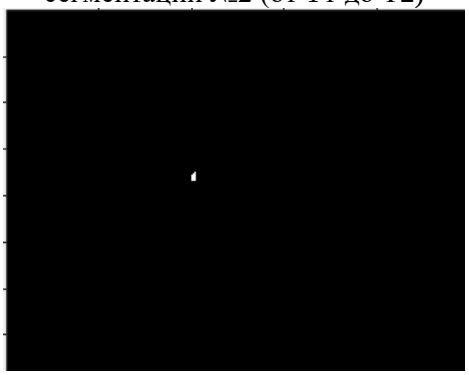


Рисунок 8. Область пороговой сегментации №4 (от T3 до 255)

На следующем этапе осуществляется поиск объекта интереса среди всех выделенных областей на изображении, полученном в результате пороговой сегментации. Для этого пользователь задаёт априорные параметры, которые определяют характеристики искомых объектов. Эти параметры могут включать количество объектов (один или несколько), форму (соотношение сторон), размер и степень заполненности. Алгоритм поиска включает несколько шагов:

1. Программа находит и подсчитывает связанные компоненты в бинарном изображении.
2. Извлекаются параметры каждого компонента, такие как наименьший прямоугольник, содержащий область, его ширина и высота, коэффициент заполненности.
3. Все сегменты, площадь которых меньше заданной, удаляются.
4. Удаляются сегменты с заполненностью меньше заданного коэффициента.
5. Удаляются сегменты, не соответствующие заданному соотношению сторон.
6. Если пользователь задал поиск одного объекта, то для выбора наилучшего варианта используется метод максимизации суммы квадратов основных параметров.
7. Программа аннотирует выбранный объект ограничивающей рамкой.

В случае поиска нескольких объектов пункт 6 опускается, и все объекты, прошедшие предыдущие этапы фильтрации, считаются удовлетворительно соответствующими заданным параметрам.

V. Сравнение с классическими методами и расчёт характеристик

Оценка эффективности алгоритма требует комплексного подхода, учитывающего не только истинно положительные и ложноположительные результаты. Важно понимать, что простое сравнение по значениям может привести к некорректным выводам о качестве сегментации изображений.

Accuracy (точность, правильность) показывает долю правильных классификаций от общего числа случаев. Precision (точность) показывает, какая доля положительных предсказаний алгоритма действительно верна. Recall (полнота) показывает, какая доля действительно положительных объектов была корректно идентифицирована алгоритмом. F1-метрика — гармоническое среднее между точностью (precision) и полнотой (recall). Это сбалансированная метрика для оценки моделей классификации, особенно в условиях несбалансированных данных, где один класс встречается гораздо чаще, чем другой.

Для эксперимента использовался набор малоконтрастных изображений, размером 1024×1024 пикселей [1]. Были внесены следующие параметры для сегментации и поиска объекта интереса:

1. Количество областей сегментации (количество порогов) — 4
2. Минимальная площадь объекта — 1000
3. Минимальный коэффициент заполненности — 0,65
4. Минимальный коэффициент соотношения сторон выделенного объекта — 0,8

Результаты рассчитанных метрик были усреднены (табл. 1) и показаны на диаграмме на рис. 9.

Таблица 1. Усреднённые значения метрик

Метрика	Генетический алгоритм	Метод Отцу	Метод Капура	Метод Тсаллиса
Accuracy	0,928	0,832	0,829	0,799
Precision	0,662	0,74	0,767	0,584
Recall	0,689	0,478	0,162	0,581
F1-score	0,596	0,43	0,209	0,412

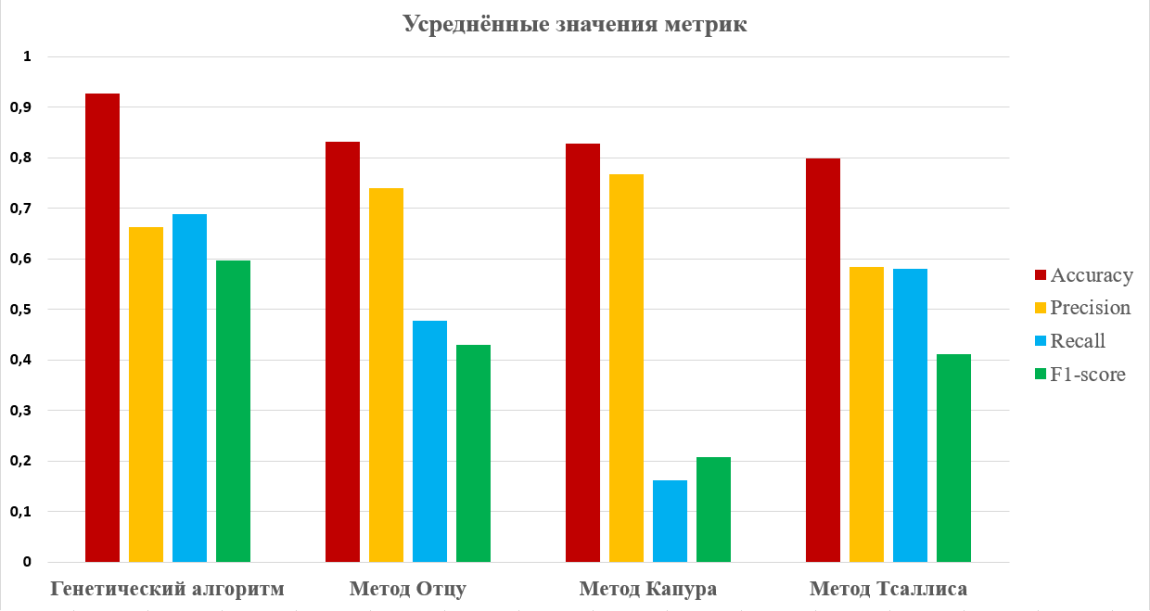


Рисунок 9. Диаграмма усреднённых экспериментальных значений метрик

На основе полученных данных можно сделать вывод, что генетический алгоритм обеспечивает повышение качества сегментации по метрикам accuracy на 11,53% и по F1 на 38,62% по сравнению с методом Отцу. Значения метрик precision и recall практически равны, из чего можно судить о том, что алгоритм демонстрирует сбалансированную работу между точностью и полнотой: хорошо избегает ложных срабатываний и эффективно обнаруживает большинство истинных положительных объектов.

VI. Заключение

В ходе проведенного исследования был реализован алгоритм сегментации, проведены сравнительные исследования с классическими методами пороговой сегментации. Рассчитаны некоторые метрики, характеризующие качество сегментации.

Использование генетического алгоритма (GA) для сегментации показало значительное улучшение в выделении контрастных объектов интереса на тестовой выборке изображений. GA эффективно идентифицирует и разделяет различные компоненты изображения, включая шум, объекты интереса и яркие блики, что позволяет целенаправленно выделять нужную часть изображения, содержащую требуемый объект интереса.

В будущем планируется улучшить алгоритмы сегментации за счет увеличения выборки, добавления постобработки результатов и расширения функционала для отслеживания объектов на последовательности кадров.

Литература

1. Я.А. Буянкина, П.В. Бабаян Пороговая сегментация изображений на основе генетического алгоритма, ГрафиКон 2025: материалы 35-й Международной конференции по компьютерной графике и машинному зрению/ Поволжский государственный технологический университет; редкол: А. А. Роженцов (отв. ред.) [и др.] – Йошкар-Ола: Поволжский государственный технологический университет, 2025. – 1095 с.
2. Nobuyuki Otsu A Threshold Selection Method from Gray-Level Histograms — IEEE Transactions on Systems, Man, and Cybernetics, 1979, vol.1(9), pp.62-66
3. Amelia Carolina Sparavigna. On the Role of Tsallis Entropy in Image Processing — International Scientific Research Journal, 2015, vol.1(6), pp.16-24
4. P.D. Sathya, R. Kayalvizhi PSO-Based Tsallis Thresholding Selection Procedure for Image Segmentation — In-ternational Journal of Computer Applications (0975 – 8887), vol. 5– No.4, August 2010, pp. 39-46
5. Кривошеев А.С., Кривошеева Е.А. Методы пороговой сегментации изображений: сравнительный анализ // Информатика, управление и системный анализ. – 2021. – Т. 5, №1. – С. 58-67.
6. Лавров В.В., Рындин Ю.Г., Тюрин В.С. Характеристики качества малоконтрастных цифровых изображений неизвестного объекта //Евразийский Союз Ученых. – 2019. – Т.1, №3. – С. 40-43
7. Omar Banimelhem, Yahya Ahmed Yahya Multi-Thresholding Image Segmentation Using Genetic Algorithm — Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV), 2011
8. Yu, W., Huang, M., Zhu, D., Li, X. (2011). A Method of Image Segmentation Based on Improved Adaptive Genetic Algorithm. — Foundations of Intelligent Systems. Advances in Intelligent and Soft Computing, vol 122. Springer, Berlin, Heidelberg

SEARCHING FOR AN OBJECT OF INTEREST IN LOW-CONTRAST IMAGES USING THRESHOLD SEGMENTATION BASED ON A GENETIC ALGORITHM

stud. Buyankina Y.A., Ph.D.Eng. Assoc. Prof. Babayan P.V.

Ryazan State Radio Engineering University named after V.F. Utkin

The actual task of image segmentation, especially in cases where the histogram of an object has several peaks (multimodality), is the subject of this study. To effectively solve this problem, a genetic algorithm was selected and implemented. The paper presents a detailed implementation of the algorithm, as well as experimental studies using various images. The analysis of the results obtained, including the calculation of key characteristics, confirms the superiority of the proposed genetic approach over traditional segmentation methods such as the Otsu method and the entropy methods of Kapur and Tsallis.

СИСТЕМАТИЧЕСКИЕ ОШИБКИ И ИХ КОРРЕКЦИЯ В СИСТЕМАХ ИЗМЕРЕНИЯ ИМПЕДАНСА ЧЕЛОВЕКА

н.с., к.ф.-м.н. Туйкин Т.С.

ФГБУН ИРЭ им. В.А. Котельникова РАН

В статье проводится анализ систематических ошибок измерения импеданса тела человека, возникающих в реальных системах из разных источников: от паразитных ёмкостей и кабелей, до мультиплексоров и фазовых детекторов. Представленная оценка степени их влияния на получаемые результаты, выполненная на основе моделирования в SPICE среде эквивалентной схемы с использованием типовых значений параметров электронных элементов, паразитных связей, импеданса тканей человека. Рассмотрены методы минимизации, стабилизации и коррекции возникающих систематических ошибок.

Введение: импеданс и как его измерять

Биоимпедансный анализ состава тела человека — это метод исследования, который позволяет оценить значения параметров состава тела и получить информацию о его состоянии. Данный метод отличается относительной простотой и быстротой проведения, при этом получаемые данные хорошо коррелируют с результатами более сложных и дорогостоящих исследований, таких как МРТ. Для специалистов в области диетологии биоимпедансный анализ предоставляет информацию о динамике жировой и мышечной ткани, что помогает в разработке индивидуальных программ питания и коррекции веса. В реаниматологии данный метод позволяет определить распределение внеклеточной и внутриклеточной жидкости, что критически важно для контроля динамики лечения пациентов в тяжёлом состоянии. Спортивные врачи могут использовать результаты биоимпедансного анализа для оценки уровня питания, интенсивности физических нагрузок и, главное, для выведения спортсменов на пик формы к соревнованиям с целью демонстрации максимальных спортивных результатов.

Для большинства практических задач проводятся два измерения на частотах 5 и 50 кГц [1] тетраполярной схемой подключения: через два электрода I+ и I- пропускается ток, на двух электродах U+ и U- измеряют падение напряжения и вычисляют импеданс тела Zb. Диагностическую ценность представляют обе компоненты импеданса модуль |Zb| и фаза Zφ, а также их соотношение на измеряемых частотах. В идеальном случае тетраполярная схема позволяет измерить импеданс непосредственно тела Zb или его исследуемого сегмента, без учета влияния контактного сопротивления кожи и электродов. Однако, на практике всегда есть конструктивные паразитные ёмкости, шунтирующие генератор тока Cg, измеритель напряжения Cv и электроды Ccab. Ccab обусловлена длинными кабелями, которые невозможно исключить при измерении импеданса человека. Поэтому в реальных системах контактный импеданс электрод-кожа Zs на практике влияет на результаты измерений. На рисунке 1 представлена модель измерения биоимпеданса, дополненная паразитными ёмкостями. При этом модель Коула для импеданса тканей [2] заменена на единичный элемент со значение суммарного импеданса ткани на заданной частоте.

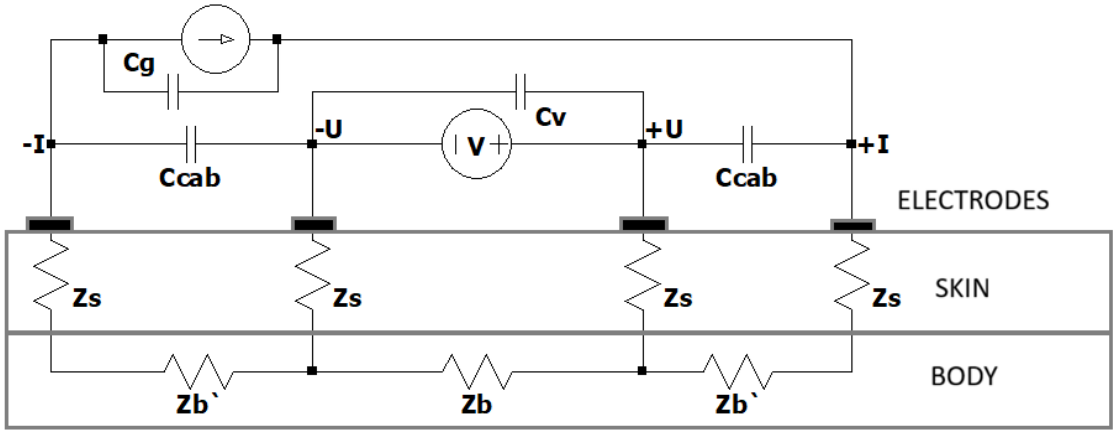


Рисунок 1 – модель измерения биоимпеданса дополненная паразитными ёмкостями

Цель работы

Провести анализ физических элементов, используемых в реализациях систем для измерения импеданса биологических объектов, на предмет выявления вносимых ими систематических ошибок в результаты измерений. Выполнить оценку их степени влияния и рассмотреть возможности компенсации.

Материалы и методы

Для анализа были использованы типовые значения биоимпеданса человека и электронных компонентов приведенные в таблице 1. Типовые значения биоимпеданса человека на основе выборки из 6-ти людей в возрасте от 26 до 73 лет, разного пола и рода деятельности. Для электронных компонентов и конструкционных элементов использованы значения из технических описаний или непосредственно измеренные. Оценка степени влияния выполнена моделированием схемы измерения биоимпеданса в SPICE среде с внесением дополнительных элементов и реальных параметров: фильтров, паразитных ёмкостей, токов утечек. Рассматривали дополнительные элементы: мультиплексоры, изолирующие конденсаторы, токоограничивающие резисторы, защитные диоды, выравнивающие потенциал резисторы. За пределами данной работы оставлено влияние электрогальванических процессов, шумы измерений.

Таблица 1. Типовые значения параметров

Параметр	**Типовое значение**
Контактный импеданс кожи Для электрода 22x34 мм	5 кГц |Z| = 0,6-3 кОм, Zφ = 20°-35° 50 кГц |Z| = 0,1-0,5 кОм, Zφ = 60°-80°
Импеданс тела	|Z| = 200-700 Ом, Zφ = 4-10°
Типовой полный импеданс тела и кожи	1 кОм – 10 кОм, Zφ = 50°-70°
Параметры кабелей	30-100 пФ/м, до 0,1 Ом/м, до 1 мкГн/м
Усилители	10 МОм, 1 нА, 5 МГц
Ограничивающие ток резисторы	1-10 кОм
Изолирующие ёмкости	15-1000 нФ
Задающие потенциал резисторы	1-10 МОм
Диоды антистатической защиты	0,5-5 пФ
Мультиплексоры	20 Ом, 40 пФ

Результаты

Для клинических приложений требуется точность измерения 1% для модуля |Z| и 0,1° для фазы Zφ импеданса биологического объекта. Фаза наиболее чувствительна к вносимым ошибкам, поэтому степень влияния будем оценивать именно по ней, а не модулю импеданса. В таблице 2 приведено максимальное влияние факторов при типовых значениях параметров.

Таблица 2. Оценка типовых ошибок измерения фазы импеданса на частоте 50 кГц

Фактор	**Влияние**
Фильтры на частоте среза 3dB	До 45°
Изолирующие ёмкости и токоограничивающие резисторы	До 20°
Паразитная емкость источника тока 100pF	До 5°
Паразитная емкость измерителя напряжения 100pF	До 2°
Погрешность калибровочного резистора 10кОм	До 1°
Паразитная ёмкость калибровочного резистора 5пФ	До 1°
Ошибки номиналов и калибровочных значений 1%	До 1°
Мультиплексор 8-каналов 20 Ом, 40 пФ	До 1°
Сдвоенный кабель 2м 65пФ	До 0,5°
Положение электродов ±2 см	До 0,4°
Квадратурный демодулятор	До 0,2°
Задающие постоянный потенциал резисторы 10 МОм	До 0,1°

Обсуждение

Для удобства восприятия сгруппируем факторы, влияющие на систематические ошибки по типам источников погрешностей, и рассмотрим подробнее механизмы влияния и методы компенсации.

Группа 1: коммутационные ошибки – паразитные параметры, кабели, кожа и электроды

Типовое значение модуля полного импеданса тела с кожей до 10 кОм и фазовый угол порядка 45°. Поэтому на частоте 50 кГц индуктивность кабелей, трасс печатных плат, выводов микросхем и их омическое сопротивление имеют суммарный импеданс менее 0,3 Ом и их можно не учитывать, поскольку вносимые ошибки менее 0,1°. Вклад паразитных ёмкостей более значительный и требует тщательного учёта. К примеру, незначительная паразитная ёмкость генератора в 10 пФ на частоте 50 кГц имеет импеданс 418 кОм и шунтирует весь измеряемый объект, что может давать ошибку измерения фазы порядка 2°. Вместо источника тока можно использовать источник напряжения, на который не влияет шунтирующая емкость. У него шире динамический диапазон по нагрузке, проще конструкция и, зачастую, лучше линейность. Данное влияние так же можно убрать непосредственным измерением фактического тока, протекающего через объект измерения. Такой подход в комбинации с доступными сейчас вычислительными мощностями даже на уровне микропроцессора делает измерения точнее.

Паразитная ёмкость измерителя напряжения шунтирует только импеданс тела Zb, поэтому её влияние меньше и может достигать единиц градусов. При проведении измерений следует избегать дополнительных паразитных ёмкостей, шунтирующих измеряемый объект и измеритель напряжения. В частности, не располагать пациента рядом и на проводящих поверхностях, избегать контактов частей тела, влажной одежды, следить, чтобы провода не лежали на металлических конструкциях.

К влиянию паразитных ёмкостей кабелей, шунтирующих электроды, есть два подхода: минимизировать и компенсировать. Для минимизации используют кабели с активными экранами и пространственное разнесение. Компенсацию можно осуществить аналитически, вычисляя по известным значениям паразитных параметров, или через матрицы калибровочных коэффициентов. Практическое применение находит комбинированный подход: минимизировать паразитную ёмкость между потенциальными электродами и далее, зная параметры, компенсировать и/или калибровать систему [3]. Повторяемость измерений важнее систематических ошибок, которые можно компенсировать. На практике для стабильности параметров кабелей используют стационарные измерители без подвижных внешних кабелей или сдвоенные кабели для двух соседних электродов: токового и потенциального. Гибкий сдвоенный кабель имеет меньшую погонную ёмкость и лучше сохраняет её значения при изменениях положения в пространстве, нежели коаксиальный кабель. Типовая емкость сдвоенного кабеля длинной 185 см 65 пФ на 50 кГц имеет импеданс 49 кОм, что при типовом импедансе тела человека с кожей до 10 кОм может дать ошибку измерения фазы импеданса до 0,5°.

Контактный импеданс, включающий в себя кожу, электрод и их контакт имеет значительно большую ёмкостную составляющую, чем импеданс тела. Его значение зависит от площади электродов, качества их прижима/контакта и локального состояния кожи. У одного человека на разных участках кожи отличие может достигать десятки процентов. Для практических применений с точки зрения гигиены и повторяемости контакта удобно использовать адгезивные одноразовые электроды, ошибки от их размещения составляют порядка 0,2°.

Попробуем упростить модель для понимания процессов и макета калибровочного стенда. Использование измерителей напряжения с высоким входным сопротивлением и током 10^{-9}-10^{-12} при зондирующем токе 10^{-3}-10^{-4} А позволяет не учитывать влияние контактного импеданса Zs под электродами U+ и U-, измеряющими напряжение. Учитывая, что импеданс участка тела между токовым и потенциальным электродами Zb` составляет единицы процентов от импеданса тела Zb и на 2-3 порядка меньше контактного импеданса Zs. На рисунке 2 представлена упрощенная модель.

Группа 2: ошибки номиналов и калибровочных значений

Системы измерения биоимпеданса обычно требуют калибровки из-за нелинейности источников тока и влияния контактного импеданса кожи на результаты измерений. При

использовании внешних калибровочных средств, принципиально важно знать их фактические значения параметров с точностью выше 0,1%, включая паразитные ёмкости. Использование переключаемых значений номиналов в калибровочном стенде может давать паразитную ёмкость десятки пФ и даже зависеть от коммутации выбранных номиналов. Отдельное внимание стоит уделить деградации ёмкостей со временем.

Наличие в составе системы внутреннего калибровочного элемента позволяет измерять параметры внутренних интегральных элементов, текущие значения параметров, компенсировать их изменения и адекватно оценивать работоспособность системы в целом. Целесообразно использовать высокоточный резистор с допуском 0,1% соразмерный измеряемому биоимпедансу до 10 кОм. Его паразитную ёмкость необходимо учитывать, поскольку несколько лишних пФ, на частоте 50 кГц внесут ошибку в измерение фазы порядка 1°.

Отличие фактических значений от номиналов и температурный дрейф вносит искажения в первую очередь через работу усилителей и фильтров см. ниже. Измерения биоимпеданса требуют оголённой кожи, поэтому обычно происходят в стабильных комнатных условиях, влияние температурного дрейфа возможно только при использовании системы сразу после транспортировки или хранения в иных температурных условиях, а также в процессе прогрева внутренних структур при включении и изменении режима работы.

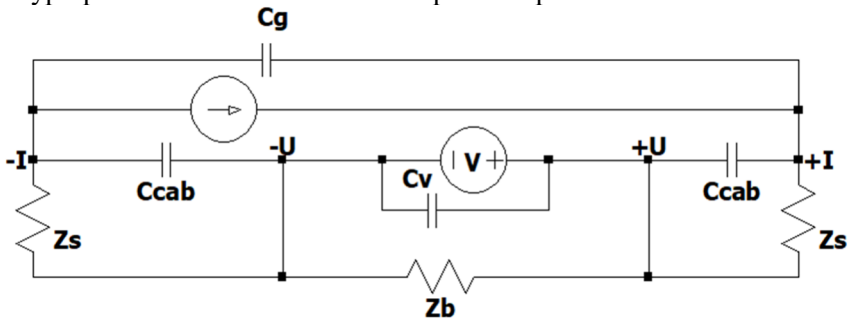


Рисунок 2 – упрощенная модель измерения импеданса человека для калибровочного стенда

Группа 3: ошибки фильтрации

Обычно системы измерения биоимпеданса используют активные фильтры/повторители, снимающие сигнал и выделяющие, рабочий диапазона частот с целью улучшения сигнал/шум и защиты от помех, наведенных на длинные измерительные кабели. Они вносят заметный поворот фазы сигнала. Фильтра первого порядка на частоте среза 3 дБ поворачивает фазу сигнала на 45°, при меньших частотах сдвиг пропорционален частоте. Реальный усилитель-повторитель тоже вносит сдвиг фазы из-за конечности коэффициента усиления и граничной частоты.

Измерение фактического тока и напряжения помогает минимизировать ошибки фильтров. Однако, при использовании трансимпедансного усилителя см рис. 3 ток определяется через падение напряжения Vi на резисторе обратной связи Zt. Фактический ток протекает через объект в другую сторону и его вычисление требует разворота на 180°. В этом случае происходит суммирование влияния фильтров на ошибки измерения фазы тока и напряжения.

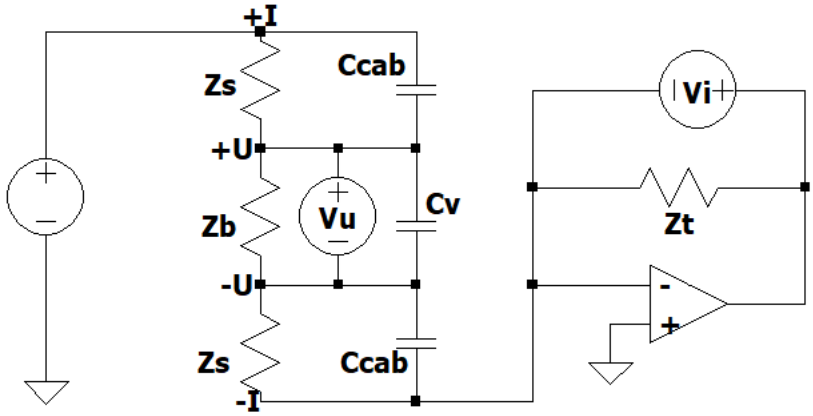


Рисунок 3 –измерение тока трансимпедансным усилителем

Группа 4: ошибки от дополнительных элементов, мультиплексоров и токов утечек

Для минимизации влияния электрогальванических процессов кожи исключают протекание постоянного тока через пациента последовательным подключением высокостабильных изолирующих ёмкостей номиналом десятки-сотни нФ. Часто используют токоограничивающие резисторы в единицы кОм для ограничения тока в целом. Влияние таких цепочек значительно и может достигать, как и у фильтров, десятков градусов. Знание их точных значений позволяет компенсировать влияние и вычислить импеданс тела человека.

С изолирующими ёмкостями к объекту часто подключают резисторы номинала 1-10 МОм для задания постоянного потенциала и отведения токов смещения, утечек.

При наличии у системы нескольких отведений используют мультиплексоры для переключения генератора и измерителя между ними. Лучшие мультиплексоры имеют сопротивление канала десяток Ом и паразитную ёмкость десятки пФ в зависимости от количества каналов. Важный момент для компенсации – учёт состояния мультиплексора, поскольку паразитная ёмкость при отключении и подключении отличается в разы.

Использование защитных диодов предупреждает повреждения входных каскадов от статического электричества и вносит дополнительно 0,5-5пФ.

Основное влияние этих факторов в виде увеличения токов утечек и разнице между измеренным/номинальным током и фактическим током, протекающим непосредственно через тело. Ёмкости дают больший вклад, чем резисторы даже при одинаковых значениях модуля импеданса из-за большей разницы фазового угла с фазой тела. На частоте 50 кГц каждые 10 пФ паразитной ёмкости на землю дают ошибку до 0,1°, каждый 1 МОм утечки до 0,05°.

Группа 5: ошибки фазового детектирования

Система, измеряющая амплитуду и фазу посредством квадратурного демодулятора, выполненного на полупроводниковых интегральных ключах, имеет наиболее простую и стабильную конструкцию. Наибольшие искажения при работе ключа в демодуляторе возникают из инжекции заряда при его переключении и мёртвого времени, когда ключ переключается и не подключен ни к одному из входов. Типовая инжекция ключа порядка 1 пКл на частоте 50 кГц эквивалентна ошибке в измерении тока на 50 нА. Мёртвое время ключа меняет окно выборки и при типовом значении 10 нс на частоте 50 кГц создаёт ошибку тоже порядка 0,1°.

Использование для измерения фазы и амплитуды быстрых АЦП и последующих преобразований Фурье значительно улучшает качество измерений. Благодаря развитию микроэлектроники и вычислительных мощностей такое решения стало широко доступно даже на уровне микропроцессорной измерительной системы без привлечения ЭВМ. Также их несомненным достоинством является стабильность во времени и независимость от разброса номиналов, их изменения от времени и температуры. Для биоимпедансных измерений не требуется высокая скорость измерения, что позволяет дополнительно поднимать соотношение сигнал/шум путём увеличения время усреднения. Однако, усреднять более 50 мс нецелесообразно, так как это вносит шумы от биологических процессов сердцебиения и дыхания. При наличии внутреннего эталонного резистора и процедуры самокалибровки можно компенсировать искажения амплитуды сигнала.

Группа 6: ошибки квантования частот

Непосредственно ошибки выставления зондирующей частоты при использовании современных кварцевых генераторов незначительны. Для зондирующей частоты 50 кГц ошибка измерения фазы в 1° может быть получена неточностью тактирования частоты 18 МГц. Такое возможно вследствие разной кратности тактов, требуемых процессору для исполнения команд, из-за прерываний или непосредственно ошибок тактирования между элементами, внутренними шинами данных. Например, при запуске по последовательному интерфейсу наличие не сбрасываемых делителей частоты в составе элементов может внести значительные и даже случайные ошибки в измерение фазы. При использовании различных источников частот стоит уделить особое внимание этому типу ошибок.

Выводы

Наибольшие ошибки измерения значений биоимпеданса вносят фильтрующие каскады и нестабильность их элементов от времени, температуры, отличия номинальных параметров от фактических значений. Именно поэтому большинство систем требуют калибровки на внешних стендах. При калибровке принципиально важно учитывать паразитные ёмкости и знать фактические величины с точностью выше 0,1%, при этом можно использовать упрощённую схему для создания калибровочных стендов.

Вторым по степени влияния оказываются паразитные ёмкости, шунтирующие генератор тока и тело/измеритель напряжения. При измерениях главное стабильность параметров, поскольку повторяемые систематические ошибки можно компенсировать. Использование сдвоенных кабелей для подключения соседних токового и потенциального электродов является хорошим решением со стабильным изменением фазы вне зависимости от положения кабеля.

Дополнительные элементы добавляют токи утечки, чем увеличивают ошибки измерения тока, или меняют значения измеряемого импеданса при подключении к нему последовательно или параллельно. В обоих случаях можно вычислить правильные значения биоимпеданса.

Грамотная структура с минимальным, а главное, стабильным шунтированием измерителя напряжения позволяет исключить влияние контактного импеданса электродов и кожи на получаемые результаты измерений.

Быстрые АЦП и преобразование Фурье стало доступнее и позволяют поднять точность, стабильность измерений, а в комбинации с измерением тока и эталонного калибровочного резистора даёт возможность избежать влияния изменений номиналов фильтрующих элементов, и их деградации со временем.

Большинство влияний имеет линейный характер от частоты и в диапазоне возможных значений импеданса тела и электродов можно ограничится аналитической компенсаций фазы от частоты и паразитных ёмкостей или даже линейной компенсацией фазы от частоты.

Современная элементная база и вычислительные мощности позволяют создавать биоимпедансные системы, которые с использованием аналитической компенсации и/или табличных поправок обеспечивают точность измерения выше 0,1° и на первый план выходят методологические ошибки разности измерения биоимпеданса как-то: размер и тип используемых электродов, места их приложений, положение пациента при измерении.

Литература

1. Bioimpedance and Bioelectricity Basics, Sverre Grimnes, Ørjan G. Martinsen, 3rd Edition, 2014, ISBN 978-0-12-411470-8
2. Kyle UG, Bosaeus I, De Lorenzo AD e.t.c.; Composition of the ESPEN Working Group. Bioelectrical impedance analysis--part I: review of principles and methods. Clin Nutr. 2004 Oct;23(5):1226-43. doi: 10.1016/j.clnu.2004.06.004
3. Николаев Д.В., Смирнов А.В., Бобринская И.Г., Руднев С.Г. Биоимпедансный анализ состава тела человека. М.: Наука, 2009. 392 с. ISBN 978-5-02-036696-1

SYSTEMATIC ERRORS IN BIOIMPEDANCE MEASUREMENT

Tuykin T.S.

Kotelnikov Institute of Radioengineering and Electronics (IRE) of Russian Academy of Sciences

The article analyzes the systematic errors that occur when measuring bioimpedance in real systems. Various sources of error are spotted from cables to multiplexers. An assessment of the degree of their influence on the measured results is presented. The assessment based on modeling an equivalent circuit in SPICE software environmental through using typical values of the parameters of electronic elements, parasitic capacitance and values obtained by measuring human bioimpedance. Methods of minimizing, stabilizing, and correcting systematic errors are presented.

К ВОПРОСУ О РАЗМЕТКЕ ЭНДОСКОПИЧЕСКИХ ВИДЕОДАННЫХ ДЛЯ ОБУЧЕНИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

доц. Хрящев В.В., доц. Завьялов Д.В., Волков Д.Б., асп. Соргин Р.И.

Ярославский государственный университет им. П.Г. Демидова,
Ярославский государственный медицинский университет

В работе рассмотрена проблема подготовки и разметки эндоскопических видеоданных для обучения систем искусственного интеллекта в медицинской диагностике. Предложена методика полуавтоматической разметки в парадигме Human-in-the-Loop с использованием фундаментальной визуальной модели DINOv3, позволяющая выполнять интеллектуальный отбор наиболее информативных кадров на основе анализа их векторных представлений и оценки визуальной разнообразности. На базе клинических видеозаписей колоноскопических исследований сформирован датасет изображений патологий толстой кишки, включающий как нормальные участки слизистой, так и различные типы новообразований, который разделён на обучающую, валидационную и тестовую выборки. Предложенный подход обеспечивает сокращение трудозатрат при сохранении высокого качества аннотаций за счёт сочетания автоматизированного отбора данных и экспертной валидации, что способствует повышению эффективности обучения моделей искусственного интеллекта для задач классификации патологических изменений в эндоскопических изображениях.

Одной из ключевых проблем при разработке интеллектуальных систем поддержки принятия врачебных решений, основанных на современных методах искусственного интеллекта, остается ограниченная доступность крупных, репрезентативных и качественно размеченных медицинских наборов данных (датасетов). Формирование таких наборов данных требует значительных временных и организационных ресурсов, поскольку предполагает участие высококвалифицированных врачей-экспертов, соблюдение клинических протоколов и контроль качества произведенных аннотаций. Особенно остро эта проблема проявляется в области медицинской визуализации, где высокая вариативность анатомических структур, наличие артефактов и редкость отдельных патологий затрудняют создание унифицированных обучающих выборок. В связи с этим всё большее распространение получают методы полуавтоматической разметки в парадигме Human-in-the-Loop, позволяющие объединить вычислительную эффективность машинных алгоритмов предварительной аннотации и клиническую точность экспертной валидации. Такой подход обеспечивает существенное снижение трудозатрат и одновременно поддерживает высокий уровень достоверности данных, критически важный для медицинских приложений [1].

Современное развитие систем искусственного интеллекта в медицине характеризуется переходом от специализированных моделей глубокого обучения, ориентированных на узкие задачи, к универсальным фундаментальным визуальным моделям (Foundation Vision Models), способным формировать обобщённые представления для визуальной информации. Данный сдвиг обусловлен, прежде всего, недостатком масштабных медицинских датасетов с детальной разметкой, стоимость создания которых значительно выше по сравнению с другими областями компьютерного зрения. Фундаментальные модели позволяют частично преодолеть это ограничение за счёт предварительного обучения на больших объемах разнородных данных с использованием самообучающихся подходов, минимизируя необходимость в полномасштабной экспертной аннотации на этапе адаптации к конкретным клиническим задачам [2-4].

Фундаментальная визуальная модель, обученная с применением самообучающихся методов, формирует универсальное пространство признаков, пригодное для широкого спектра задач медицинского анализа изображений, включая классификацию, сегментацию, детектирование патологических изменений и выбор информативных кадров из видеопоследовательностей. Ключевыми характеристиками таких моделей являются масштабируемость, переносимость знаний между доменами, способность работать с частично размеченными данными и высокая адаптивность, что во многом стало возможным благодаря развитию архитектур трансформеров и механизмов внимания. В контексте медицинских данных это открывает возможности для более эффективного использования неаннотированных клинических архивов, существенно расширяя потенциал обучения моделей.

Примером современной фундаментальной модели является DINOv3, обучаемая в полностью самоконтролируемом режиме на больших массивах изображений без прямого использования

меток классов. В процессе обучения модель выявляет устойчивые визуальные паттерны и формирует информативные векторные представления, сохраняющие структурные и текстурные особенности изображений. Для задач анализа эндоскопических видеоданных важными преимуществами DINOv3 являются способность эффективно извлекать признаки из неразмеченных медицинских данных, высокая чувствительность к тонким текстурным изменениям слизистой и устойчивость к типичным для эндоскопии визуальным артефактам, таким как блики, размытие и изменение освещения [5-9].

В настоящей работе предварительно обученная модель DINOv3 используется для оптимизации процесса разметки эндоскопических видеоданных посредством этапа интеллектуального «прореживания» кадров. Каждый кадр преобразуется в высокоразмерное векторное представление (embedding), после чего между дескрипторами вычисляется косинусное расстояние, позволяющее оценить степень визуального различия изображений. На основе данного критерия отбираются наиболее информативные и непохожие кадры, что снижает избыточность данных и сокращает объём материала, требующего экспертной аннотации. Предложенный подход позволяет существенно уменьшить временные затраты на подготовку датасета и одновременно повысить эффективность последующей адаптации моделей искусственного интеллекта к специфике эндоскопической визуализации, что является важным шагом на пути к созданию клинически применимых интеллектуальных систем.

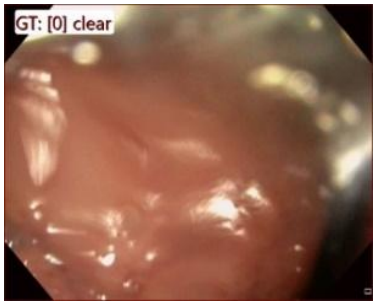
На базе Ярославской областной клинической онкологической больницы (ЯОКОБ) собраны 180 видеозаписей колоноскопических обследований. Средняя длительность каждой видеопоследовательности составила около 10 минут. Объект интереса представлен на нескольких кадрах, что улучшает обнаружение и классификацию, но требует повышенных вычислительных ресурсов и сложных алгоритмов обработки. Поэтому в рамках настоящей работы данной работе корреляция между кадрами не учитывается.

В качестве исследуемых классов патологий были использованы:

- Класс 1. Области без патологий (clear).
- Класс 2. Аденома (adenomatous).
- Класс 3. Гиперпластический полип и зубчатые новообразования (hyperplastic).
- Класс 4. Рак с поверхностной инвазией (cancer_t1).
- Класс 5. Рак с глубокой инвазией (cancer_t2).

Аденома кишечника представляет собой доброкачественное новообразование эпителиального происхождения, формирующееся из железистых клеток слизистой оболочки кишечной стенки. Гиперпластический полип кишечника относится к доброкачественным изменениям слизистой и характеризуется локальным утолщением ткани, обусловленным пролиферацией клеточных элементов без признаков неопластической трансформации. Рак с поверхностной инвазией (T1, ранний колоректальный рак) определяется как злокачественное новообразование, распространяющееся за пределы слизистой оболочки с инфильтрацией подслизистого слоя, однако без вовлечения мышечной оболочки кишечной стенки. Рак с глубокой инвазией характеризуется проникновением опухолевого процесса в мышечный слой (стадия T2) либо распространением через все слои кишечной стенки с возможным поражением серозной оболочки и прилежащих тканей, что соответствует стадиям T3–T4.

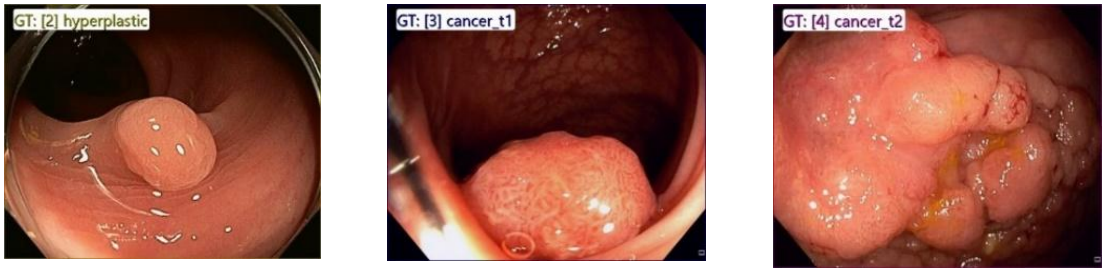
Примеры изображений с описанными выше патологиями представлены на рисунке 1.



а)



б)



в) г) д)

Рисунок 1 – Колоноскопические изображения: а) без патологии, б) с аденомой, в) с гиперпластическим полипом, г) с поверхностной инвазией, д) с глубокой инвазией

Исходные видеопоследовательности были разделены на обучающую, валидационную и тестовую выборки на уровне отдельных клинических исследований, что позволило исключить пересечение данных между подмножествами. Из каждой видеозаписи выделялись два класса изображений: кадры, содержащие патологические изменения, и кадры без признаков патологии. Отбор наиболее информативных и репрезентативных изображений для формирования датасета осуществлялся с использованием фундаментальной визуальной модели DINOv3, обеспечивающей оценку визуальной разнообразности кадров. Сводные количественные характеристики сформированного набора данных представлены в таблице 1.

Таблица 1 – Набор изображений для обнаружений патологий толстой кишки

Выборка / Класс	Обучающая, шт.	Тестовая, шт.	Валидационная, шт.
Класс 1	26023	400	2921
Класс 2	40439	1300	4561
Класс 3	33507	1400	3617
Класс 4	17307	1900	1971
Класс 5	9559	230	1029
Всего	126835	5230	14099

В результате в рамках настоящего исследования предложена методика полуавтоматической разметки медицинских изображений, основанная на применении фундаментальной визуальной модели DINOv3. Реализация данного подхода позволила существенно снизить трудоемкость формирования датасета за счёт сочетания автоматизированного отбора уникальных кадров алгоритмом и последующей экспертной валидации полученных изображений специалистами.

Доклад подготовлен в рамках Государственного задания на осуществление научных исследований и разработок ФГБОУ ВО ЯГМУ Минздрава России на 2026 г. по теме «Система поддержки принятия врачебного решения при дифференцировке глубины инвазии новообразований толстой кишки на основе искусственного интеллекта» (№720000Ф.99.1.БН62АБ20000).

Литература

1. Карпов, О.Э. Информационные технологии, вычислительные системы и искусственный интеллект в медицине / Карпов О.Э., Храмов А.Е. – М.: ДПК Пресс. 2022.
2. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение // ДМК-Пресс. 2017. 652 с.
3. Kaminski M.F. et al. Quality indicators for colonoscopy and the risk of interval cancer // New England Journal of Medicine. – 2010. – vol. 362(19). – P. 1795–1803.
4. Kaiming H., Xiangyu Zh., Shaoqing R., Jian S. Deep Residual Learning for Image Recognition // Computer Vision and Pattern Recognition (cs.CV). arXiv preprint arXiv:1512.03385 [cs.CV], 2015.

5. Liu, Z. A ConvNet for the 2020s [электронный ресурс] / Liu Z., Mao H., Wu C. – Режим доступа: https://arxiv.org/abs/2201.03545.
6. Dani A. P., Salehi I., Rotithor G., Trombetta D., Ravichandar H. Human-in-the-loop Robot Control for Human-Robot Collaboration: Human Intention Estimation and Safe Trajectory Tracking Control for Collaborative Tasks // arXiv preprint arXiv. – 2024. – March 27. – Электрон. дан. – URL: https://arxiv.org/abs/ (дата обращения: 06.11.2025).
7. Aad G. et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC // arxiv.org. – 2025. – URL: https://arxiv.org/html/2508.10104v1 (дата обращения: 06.11.2025).
8. А. Д. Борсук, Е. Г. Малаева Колоректальные аденомы // Проблемы здоровья и экологии. 2011. №3 (29). URL: https://cyberleninka.ru/article/n/kolorektalnye-adenomy (дата обращения: 07.11.2025).
9. Sano Y., Tanaka S., Kudo S., Saito S., Matsuda T. Narrow-band imaging (NBI) magnifying endoscopic classification of colorectal tumors proposed by the Japan NBI Expert Team / Digestive Endoscopy. – 2016. – Vol. 28. – P. 526–533.

ON ENDOSCOPIC VIDEO DATA LABELS FOR TRAINING ARTIFICIAL INTELLIGENCE SYSTEMS

Docent V.V. Khryashchev, Docent D.V. Zavyalov, D.B. Volkov, R.I. Sorgin

P.G. Demidov Yaroslavl State University, Yaroslavl State Medical University

The paper addresses the problem of preparing and labeling endoscopic video data for training artificial intelligence systems in medical diagnostics. A semiautomated labeling approach based on the Human-in-the-Loop paradigm and the foundation vision model DINOv3 is proposed, enabling intelligent selection of the most informative frames through analysis of high-dimensional embeddings and visual diversity estimation. Using clinical colonoscopy video recordings, a dataset of colon pathology images was created, including both normal mucosal areas and various types of neoplastic lesions, and divided into training, validation, and test subsets. The proposed approach significantly reduces annotation effort while maintaining high labeling quality through a combination of automated frame selection and expert validation, thereby improving the efficiency of training AI models for classification of pathological findings in endoscopic imagery.

ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ СИГНАЛА ЭКГ И ИХ ПРИМЕНЕНИЕ В ЗАДАЧЕ ОБНАРУЖЕНИЯ ФИБРИЛЛЯЦИИ ПРЕДСЕРДИЙ

Огнев А.П.

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

Работа посвящена исследованию информационных характеристик сигнала электрокардиограммы и оценке их применения в задаче обнаружения нарушения сердечного ритма, в частности, фибрилляции предсердий. Предложен новый метод классификации, основанный на построении информационных диаграмм статистической сложности и энтропии сигнала ЭКГ. Преимуществом данного метода является возможность его применения на коротких интервалах ЭКГ. Предложенный подход также может быть применим для анализа других периодических биомедицинских сигналов.

Введение

Сердечно-сосудистые заболевания являются основной причиной смертности во всем мире, среди которых значительную часть занимают различные нарушения сердечного ритма. Среди различных форм нарушения ритма значительную долю занимают трепетание предсердий (ТП) и фибрилляция предсердий (ФП), которые значительно увеличивают риск инсульта [1].

В частности, фибрилляция предсердий является наиболее распространенной аритмией, которой подвержены до 2% населения и до 10% лиц старше 60 лет, причем эти показатели имеют тенденцию к постоянному росту.

Большинство методов обнаружения аритмии требуют достаточно длительный интервал наблюдения, так как эти методы основаны на анализе временных рядов последовательности 200 – 300 RR–интервалов кардиограммы.

В данной работе рассматриваются спектральные методы обнаружения нарушений сердечного ритма, в частности использование энтропии Шеннона [2] сигнала электрокардиограммы, а также статистической сложности сигнала, которая определяется, как произведение энтропии на дисбаланс [5,6].

Сердечный ритм в норме и патологии

К настоящему времени известно более 20 видов нарушения сердечного ритма. К опасным относятся: фибрилляция предсердий, трепетание предсердий, к особо опасным: фибрилляция желудочков, трепетания желудочков, подьем интервала ST (возможный признак инфаркта миокарда). В данной работе рассматривается фибрилляция предсердий (мерцательная аритмия), как наиболее распространенный (более 40% среди всех видов нарушений ритма) и опасный вид аритмий.



Рис. 1. Сигнал ЭКГ в норме (синусовый ритм)

В норме форма сигнала ЭКГ сохраняется, наблюдается некоторая вариабельность интервала RR, то есть незначительные отклонения мгновенной частоты сердечных сокращений (ЧСС) от среднего значения, которое зависит от текущего состояния организма. При патологии наблюдается значительные изменения формы сигнала ЭКГ (исчезновение одних зубцов и появление других), а интервал RR хаотически изменяется.

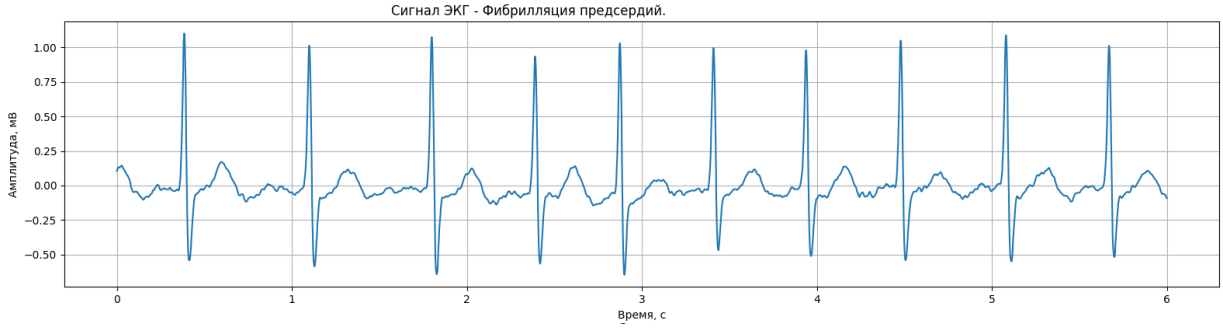


Рис. 2. Сигнал ЭКГ, имеющий признаки фибрилляции предсердий.

Фибрилляция предсердий (ФП) – нарушение ритма сердца с нерегулярными сокращениями по силе и частоте, вызванное хаотическим характером проведения волны возбуждения через миокард предсердий. Признаки ФП на кардиограмме – значительная нерегулярность RR–интервалов, а также отсутствие зубцов Р, отражающих деполяризацию предсердий. Данная патология приводит к опасным нарушениям гемодинамики, турбулентности кровотока в камерах сердца и коронарных артериях, что может привести к тромбоэмболических осложнениям, включая инсульт, поэтому, очень важно своевременно выявить наличие ФП и принять соответствующие меры.

Автокорреляция сигнала ЭКГ в окне длительностью 6 сек отражает хаотичность сигнала при фибрилляции и регулярность при нормальном синусовом ритме.

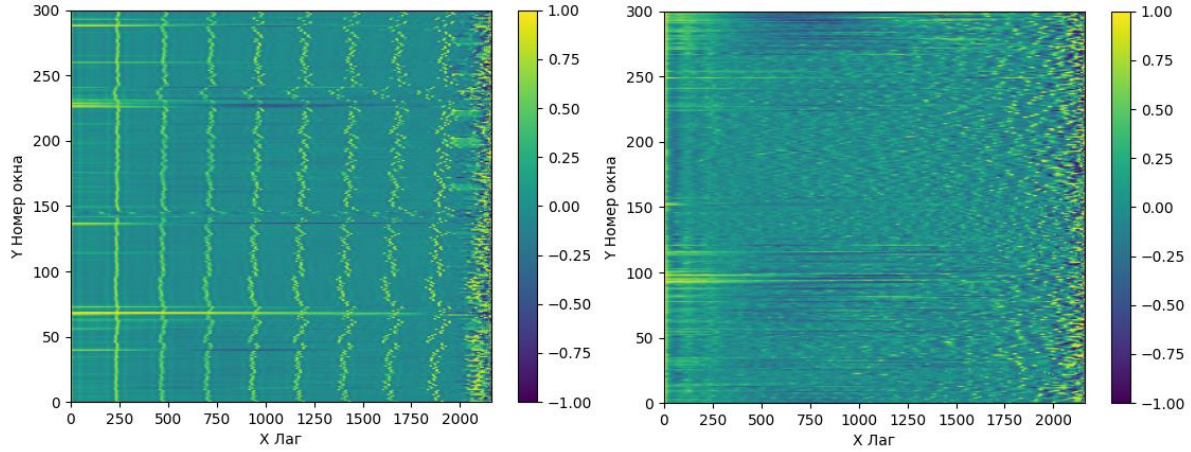


Рис. 3. Автокорреляция сигнала ЭКГ, имеющего признаки нормального синусового ритма (слева) и имеющего признаки фибрилляции предсердий (справа).

Методы оценки энтропии статистической сложности сигнала ЭКГ

Энтропия *H(X)* единичной дискретной случайной величины *X* является мерой ее средней неопределенности. Наиболее часто применяется энтропия Шеннона [3]:

$$H = -\sum_{i=0}^{n} p_i \, log_2 p_i \qquad (1)$$

где

$$p_i = \frac{x_i^2}{\sum_{i=1}^{n} x_i^2} \qquad (2)$$

– вероятность того, что сигнал содержит i-ю компоненту дискретного преобразования Фурье, выраженная как доля энергии сигнала этой компоненты (дискреты)

$$X_k = \sum_{n=1}^{N} x_n e^{-\frac{2i\pi k(n-1)}{N}} \quad (3)$$

Для определения энтропийных характеристик длительности RR-интервала сигнала ЭКГ, таких как аппроксимированная энтропия (ApEn), сэмплированная энтропии (SampEn), многомасштабная энтропия (MsEn), вейвлет–энтропия (WEn), энтропия Риньи (REn), требуются, как правило, длинные временные ряды (более 300 RR-интервалов) и большое количество данных для получения полезных результатов. [ ]

Для анализа сигнала ЭКГ на коротких интервалах предлагается использовать оконное преобразование Фурье, с длительностью окон 3, 6, 12 секунд, и метод диаграмм энтропия-сложность, когда каждое окно сигнала представлено на диаграмме в виде точки, в соответствии со значениями энтропии Шеннона и статистической сложности, равной произведению энтропии Шеннона на дисбаланс [5]:

$$C_{TV}(p) = -\frac{1}{4\ log_2 N}\left(\sum_{n=1}^{N} p_i\ log_2\ p_i\right)\left(\sum_{n=1}^{N}\left|p_i - \frac{1}{N}\right|\right)^2 \quad (4)$$

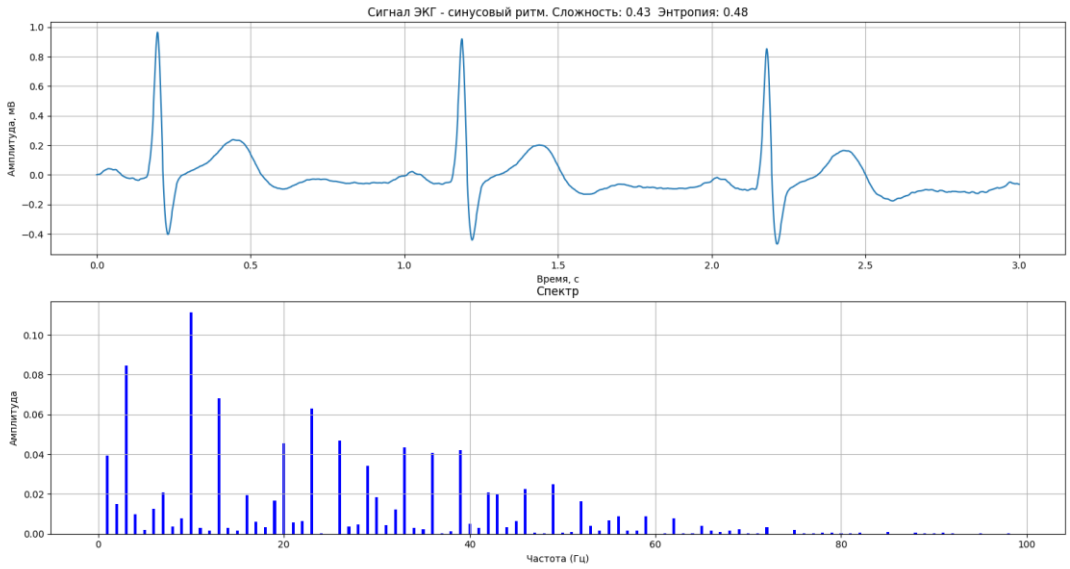


Рис. 4. Сигнал ЭКГ, имеющий признаки нормального синусового ритма, и его спектр, окно длительностью 3 сек.

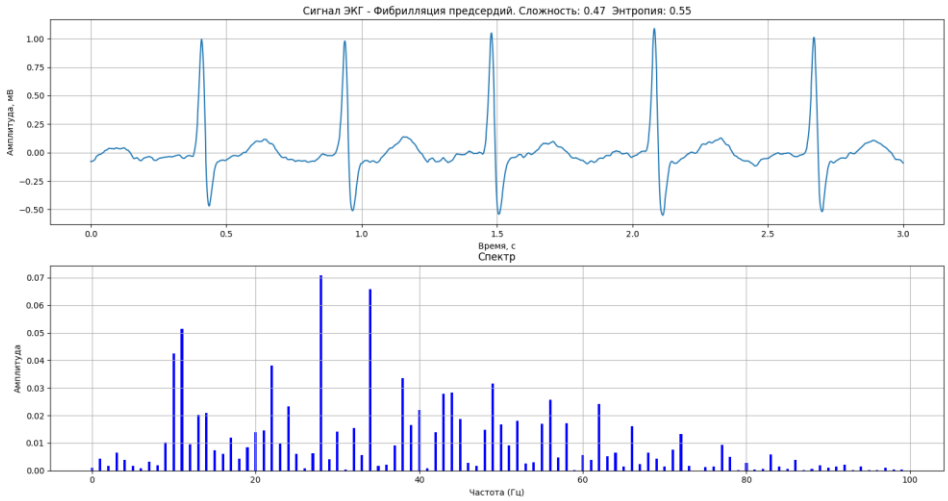


Рис. 5. Сигнал ЭКГ, имеющий признаки фибрилляции предсердий и его спектр, окно длительностью 3 сек.

Из анализа спектральной энтропии и сложности можно сделать вывод, что нормальный синусовый ритм соответствует низким значениям энтропии и сложности, оценочно $Cv < 0.45$ и $H < 0.52$. Точные границы и решающие правила классификации предлагается установить с помощью метода Random Forest.

Диаграммы энтропии и сложности

Для классификации сигналов ЭКГ в норме и патологии (в данном случае фибрилляции предсердий) применялся метод Random Forest [9], для обучения использовалась выборка 70% данных энтропии и сложности сигнала ЭКГ в окне длительностью 3 секунды без перекрытия. В качестве тестовой использовалась выборка оставшихся 30%.

В исследованиях использовался набор данных MIT-BIH afdb (PhysioNet.org) [10].

Сигнал ЭКГ при фибрилляции предсердий характеризуется высокой энтропией и сложностью, что обусловлено тем, что в этом режиме он является результатом действия совокупности эктопических источников сердечного ритма, рассогласованных по амплитуде, фазе и частоте, и поэтому имеющий хаотический характер.

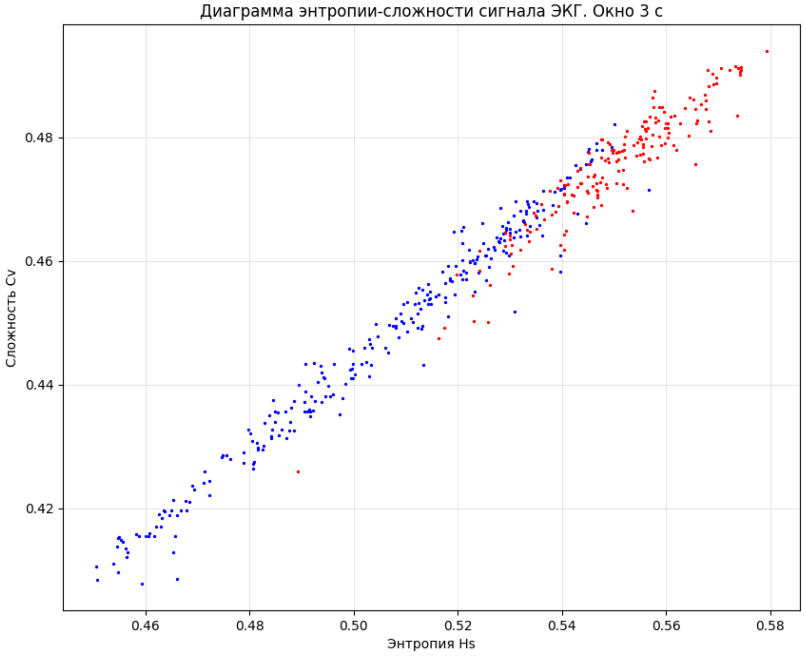


Рис. 6. Диаграмма энтропии-сложности, окна длительностью 3 сек,
· нормальный синусовый ритм, · фибрилляция предсердий,

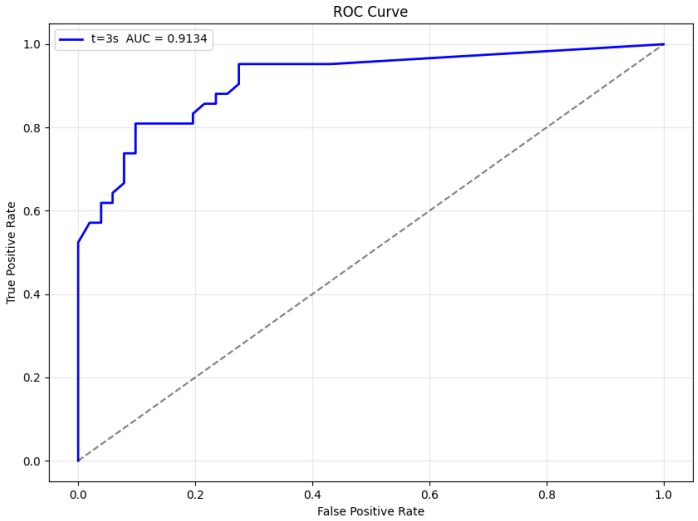


Рис. 7. ROC-кривая.

Нормальный синусовый ритм соответствует низким значениям энтропии и сложности, в то время как фибрилляция предсердий – высоким значениям энтропии и сложности. Из анализа ROC-кривой можно сделать вывод, что данный метод позволяет с достаточно высокой точностью установить наличие признаков фибрилляции предсердий на коротких интервалах ЭКГ на основе анадиза информационных характеристик сигнала. При сбалансированной чувствительности и специфичности метода его точность составляет около 90%.

Заключение

Полученные результаты демонстрируют, как можно использовать информационные методы, а именно диаграммы сложности–энтропии в целях решения задач обнаружения нарушений сердечного ритма и других периодических биомедицинских сигналов, а также классификации.

Литература

1. Martin S. et al, 2025 Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association // Circullation. – 2025. Vol. 151, № 8, – P. 379–423
2. WHO. Cardiovascular diseases (CVDs) URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (дата обращения: 21.05.2025)
3. Schuijt E, Scherr D., Plank G., Schotten U., Heijman J. Evolution in electrophysiology 100 years after Einthoven: translational and computational innovations in rhythm control of atrial fibrillation. // EP Europace, – 2025. – Vol. 27, № 1, P. euae304.
4. Shannon C.E. A Mathematical Theory of Communication. // Bell Syst. Tech. J. – 1948. Vol. 27, P. 379–423.
5. Malindzak George S. Jr. Fourier analysis of cardiovascular events. // Mathematical Biosciences. – 1970. – Vol. 7, № 3–4, P. 273–289
6. Бодин О.Н., Иванчуков А.Г., Полосин В.Г., Рахматуллов Ф.К. Энтропийно-параметрическая обработка электрокардиосигнала // Фундаментальные исследования. – 2015. – № 3, С. 23–27.
7. Галяев А.А., Бабиков В.Г., Лысенко П.В., Берлин Л.М. Новая спектральная мера сложности и её возможности по обнаружению сигналов в шуме. // Доклады Российской академии наук. Математика, информатика, процессы управлени–я. – 2023. – Т. 505, С. 46–49.
8. Бабиков В.Г., Галяев А.А. Диаграммы статистической и спектральной сложности. // Проблемы передачи информации. – 2024. – Т. 60, № 2, С. 25–35.
9. Breiman, L. Random Forests. // Machine Learning. – 2001. – Vol. 45, P. 532.
10. Moody G., Mark R. MIT-BIH Arrhythmia Database. Version: 1.0.0. URL:physionet.org/content/mitdb/1.0.0/ (дата обращения: 01.02.2026).

ECG SIGNAL INFORMATION CHARACTERISTICS AND THEIR APPLICATION TO ATRIAL FIBRILLATION DETECTION

A.P. Ognev

V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow.

This paper examines the information characteristics of electrocardiogram signals and evaluates their application in detecting cardiac arrhythmias, particularly atrial fibrillation. A new classification method is proposed based on constructing information diagrams of the statistical complexity and entropy of the ECG signal. An advantage of this method is its applicability to short ECG intervals. The proposed approach can also be applied to the analysis of other periodic biomedical signals.

РАЗРАБОТКА СИСТЕМЫ ЛАЗЕРНОГО ТРЕХМЕРНОГО СКАНИРОВАНИЯ ПРОСТРАНСТВА

проф. А. Л. Приоров, преп. В. П. Кирнос, студ. А. А. Мохов, студ. А. М. Некоркин

Ярославский государственный университет им. П. Г. Демидова

Представлена разработка мобильного лидара для трехмерного сканирования пространства, ориентированного на массовое применение. Конструкция основана на одном лазерном дальномере и системе наведения луча с вращающимся зеркалом, что позволяет формировать полусферическое облако точек при пониженной стоимости аппаратной части. Разработана модульная архитектура, включающая микроконтроллеры для сбора и предварительной обработки данных, одноплатный компьютер Raspberry Pi и веб-интерфейс на базе фреймворка Flask с визуализацией результатов в режиме реального времени.

В последние годы доступные по стоимости решения для трехмерного сканирования пространства становятся востребованными в задачах робототехники, мониторинга объектов и создания цифровых моделей помещений [1-4]. При этом многие коммерческие лидарные системы остаются дорогими и требуют применения специализированного программного обеспечения (ПО), что ограничивает их использование в образовательных и малобюджетных проектах.

Цель работы заключается в разработке компактной мобильной лидарной системы, объединяющей простую аппаратную часть и интуитивно понятный веб-интерфейс для визуализации облака точек в режиме реального времени. Новизна подхода состоит в сочетании бюджетного лазерного дальномера, механизма вращающегося зеркала и встроенного веб-сервера, обеспечивающего удалённый доступ к данным без установки специализированного ПО на рабочие станции. Работа ориентирована на создание прототипа, пригодного для дальнейшего применения в образовательных исследованиях, а также в составе мобильных робототехнических платформ (рис. 1).

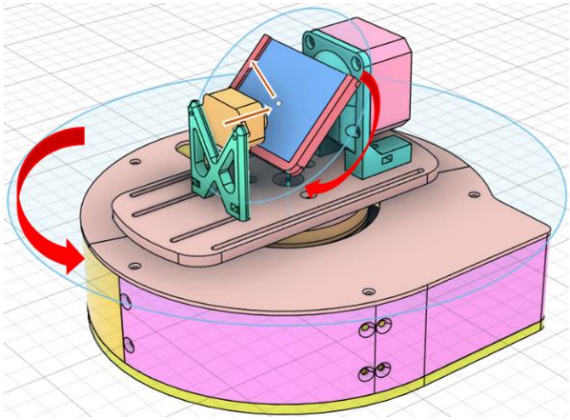


Рисунок 1 – Направления движения подвижных элементов и ход лучей лидара

Основным измерительным модулем в новой системе стал лазерный дальномер TF – Luna с рабочим диапазоном от 0,2 до 8 м. Впоследствии данный узел будет комплектоваться более точным дальномером в соответствии с требованиями заказчика. Лазерный дальномер компактен и прост в использовании, что упрощает его интеграцию в конечные устройства. Инновационность конструкции проявляется в системе позиционирования луча лидара. В основе лежит использование вращающегося зеркала, установленного под углом в 45 градусов. Оно отражает лазерный луч, обеспечивая сканирование пространства в дополнительной плоскости. Такое решение, дополненное парой высокоточных шаговых двигателей Nema 17, позволило создать систему, способную формировать полусферу пространственных данных при существенно меньшей себестоимости (рис. 2). Однако в процессе разработки возникла инженерная проблема – быстрый износ контактного кольца (Slip Ring), что приводило к потере части передаваемых данных. Это стало отправной точкой для дальнейших исследований в области надежных

контактных систем и беспроводной передачи информации на малых дистанциях (с неподвижной части системы на подвижную платформу).

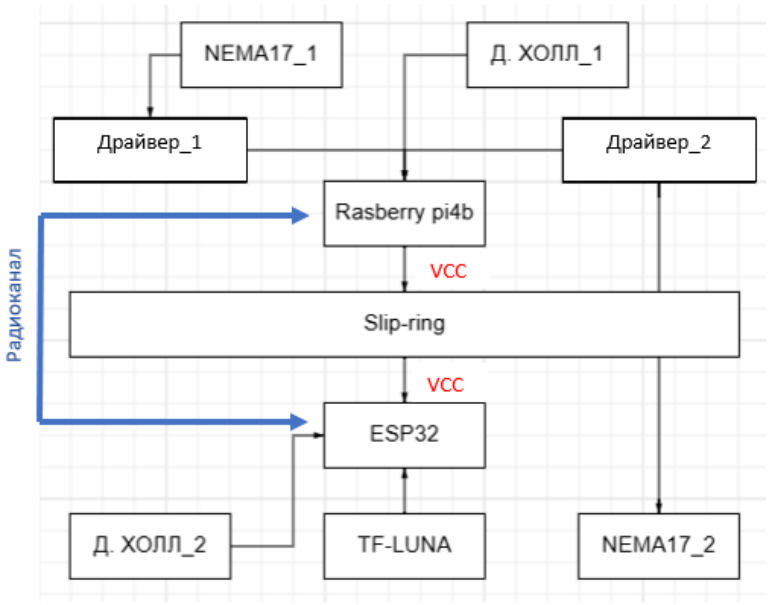


Рисунок 2 – Блок-схема оборудования лидара

Впоследствии система питания и управления была переработана. Контактное кольцо теперь служит исключительно для подачи питания, а передача данных осуществляется по беспроводному каналу. Дополнительный микроконтроллер на поворотной платформе агрегирует данные с дальномера и передает их основному – Raspberry Pi, расположенному в основании платформы, через специализированный беспроводной протокол. Это решение значительно улучшило конструкцию системы и исключило влияние механического износа контактного кольца, что важно для создания надежных устройств.

Программная часть системы изначально проектировалась как более простая альтернатива специализированным решениям. Для управления лидаром и доступа к данным разработано серверное приложение на базе фреймворка Flask, обеспечивающее работу встроенного веб-интерфейса. Подключение к системе осуществляется через обычный браузер по локальной сети, без установки дополнительного программного обеспечения на рабочие станции. Взаимодействие между сервером и клиентом реализовано с использованием технологии WebSocket, что позволяет визуализировать формируемое облако точек в режиме, близком к реальному времени, и динамически обновлять отображаемые данные (рис. 3).

Пользовательский интерфейс предоставляет базовые инструменты анализа: помимо просмотра облака точек, доступны срезы по осям X, Y и Z (рис. 4), а также масштабирование и перемещение сцены. Такой подход упрощает первичную обработку и интерпретацию результатов измерений, делает систему доступной для пользователей без опыта работы с профессиональными пакетами 3D-моделирования и позволяет исключить необходимость использования физических интерфейсов подключения (USB, Ethernet) при эксплуатации устройства.

Предполагаемая архитектура позволяет легко заменять компоненты или интегрировать дополнительные датчики, что делает систему максимально гибкой и адаптивной для будущих улучшений.

Программное обеспечение находится на этапе отладки. Завершена интеграция микроконтроллеров и проведено базовое тестирование передачи данных. В ближайшее время запланированы полевые испытания системы в различных условиях – как в помещениях, так и на открытом воздухе, а также при низкой освещенности. Эти испытания помогут оценить точность и стабильность работы системы, а также выявить возможные недостатки для доработки.

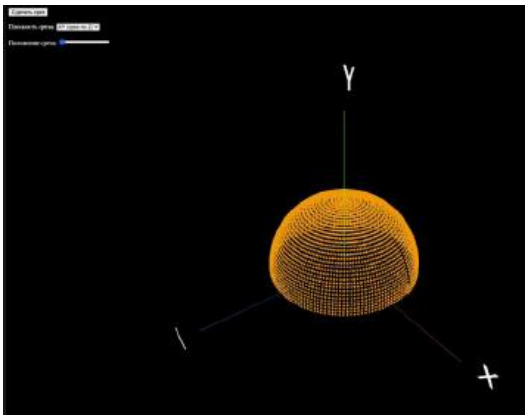
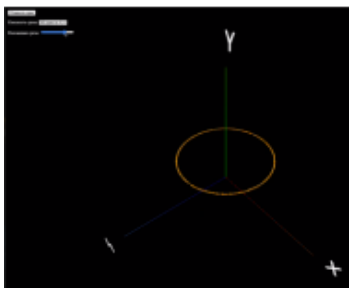
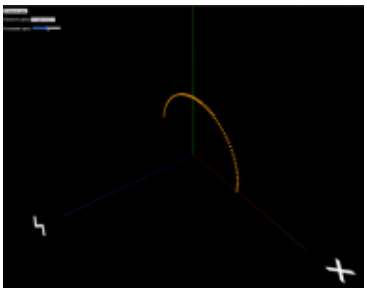
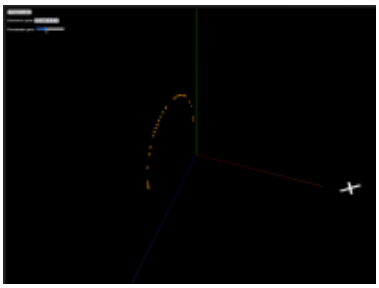


Рисунок 3 – Полусфера данных







а) б) в)

Рисунок 4 – Срезы по разным осям: а) по оси y; б) по оси z; в) по оси x

Несмотря на достижения, новая лидарная система имеет ограничения, связанные со скоростью и точностью сканирования, которые зависят от характеристик лазерного дальномера. Эти параметры могут ограничивать применение системы в некоторых специфических задачах, а также определяют направления для дальнейших исследований и улучшений.

Перспективы применения разработанной технологии широки. В ближайшей перспективе имеется возможность использования разработанной системы для сканирования жилых и промышленных помещений, навигации автономных роботизированных систем и мониторинга охраняемых территорий. Особое внимание уделяется развитию отечественной компонентной базы – идет активная работа над созданием собственных решений, которые позволят снизить зависимость от импортных комплектующих. Данный проект демонстрирует, что передовые технологии могут и должны быть более доступными.

Векторы развития:

- Замена TF-Luna на более продвинутые модели (например, Benewake TF03) для увеличения дальности сканирования,
- Реализация алгоритмов SLAM (Simultaneous Localization and Mapping) позволит использовать лидар для автономной навигации роботов [5-7],
- Переход на другие микроконтроллеры со сниженным энергопотреблением, чтобы сделать систему более мобильной.

Разработанный лидар демонстрирует жизнеспособность концепции бюджетного 3D-сканирования для образовательных и исследовательских целей. Интеграция аппаратных и программных решений позволила создать устройство, сочетающее простоту использования и функциональность. Несмотря на ограничения в точности и скорости, проект служит основой для дальнейших модификаций, включая применение в робототехнике, картографии и системах дополненной реальности. Успешная реализация беспроводного интерфейса и браузерной визуализации открывает возможности для дистанционного управления и коллективной работы с данными, что особенно актуально в условиях развития IoT-технологий. Таким образом, проделанная работа подтверждает, что даже с

использованием доступных компонентов можно создать эффективный инструмент для анализа пространства.

Литература

1. Кулешов С.В., Зайцева А.А., Аксенов А.Ю., Карпов А.А., Кипяткова И.С., Ватаманюк И.В. 3D-технологии в современных информационных системах: теория и практика. Учебно-методическое пособие. СПб: СПб ФИЦ РАН. 2021. 83 с.
2. Гура Д.А., Беспятчук Д.А., Самарин С.В., Кирюникова Н.М., Лесовая Э.Д. Технология трехмерного лазерного сканирования как средство обеспечения безопасности объектов спортивной инфраструктуры // Нанотехнологии в строительстве. 2021. № 13(4). С. 259-263.
3. Li Z., Wang Y., Chen H. et al. Towards an ultrafast 3D imaging scanning LiDAR system: a review // Journal of Innovative Optical Health Sciences. 2024. Vol. 17. N 8. pp. 1709.
4. Santos R.L., Silva M., Oliveira R. A systematic literature review of low-cost 3D mapping solutions // Journal of Network and Computer Applications. 2024. Vol. 43. N 2. pp. 321-339
5. Santos R.L., Silva M., Oliveira R. An extension of ORB-SLAM for mobile robot using LiDAR and monocular camera data for SLAM without odometry // Proc. Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO). 2024. pp. 1–8.
6. Tian W., Zhang J., Liu Y. Ground-constrained 3D LiDAR SLAM design and implementation for mobile platforms // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Archives). 2024. Vol. XLVIII-4. pp. 697–704.
7. Author A., Author B. LiDAR SLAM for mobile robot in an indoor environment // International Journal for Research in Applied Science and Engineering Technology. 2025. Vol. 13. N 5. pp. 1–7.

DEVELOPMENT OF A THREE-DIMENSIONAL LASER SPACE SCANNING SYSTEM

A. L. Priorov, A. A. Mokhov, A. M. Nekorkin, V. P. Kirnos

P. G. Demidov Yaroslavl State University

The development of a mobile lidar for three-dimensional space scanning aimed at mass use is presented. The design is based on a single laser rangefinder and a beam guidance system with a rotating mirror, which makes it possible to form a hemispherical point cloud at a reduced hardware cost. A modular architecture has been developed, including microcontrollers for data collection and preprocessing, a Raspberry Pi single-board computer and a web interface based on the Flask framework with real-time visualization of results.

ИССЛЕДОВАНИЕ СЦЕНАРИЕВ ПРИМЕНЕНИЯ РТК В УСЛОВИЯХ ЦИФРОВОГО МОДЕЛИРОВАНИЯ ПОЛУАВТОНОМНЫХ ФУНКЦИЙ

с.н.с., к.т.н. Саломатин А.А., с.н.с., к.т.н. Рожнов А.В., н.с. Журавлёва Н.Г.

Институт проблем управления им. В.А. Трапезникова РАН

Многоаспектное моделирование в условиях цифровой трансформации предъявляет новые требования при обосновании и реализации междисциплинарных результатов решения комплексных задач организации и проведения предпроектных исследований сценариев применения робототехнических комплексов (РТК) различного назначения. Цель исследования – обоснование структурной модификации «предобученных на основе теории» (TTNs) / физико-информированных (PINNs) искусственных нейронных сетей на примере командно-сигнальной системы заданной структуры применения автономных и полуавтономных функций. Совокупность взаимоувязанных на текущем этапе программно реализуемых решений включает предварительные результаты предиктивного моделирования посредством оценивания возможностей адаптации сигнально-кодовых конструкций в аспекте изучения характерных особенностей межмашинного обмена для группового управления РТК различного назначения.

Групповое управление, межмашинный обмен, многоаспектное моделирование, полуавтономные функции, предиктивное моделирование, «предобученные на основе теории» [Theory-Trained Neural Networks, TTNs] / физико-информированные [Physics-informed neural networks, PINNs] искусственные нейронные сети, предпроектные исследования, робототехнические комплексы, сквозная интеграция, сценарный подход, цифровая трансформация.

Современные тенденции интенсивного развития средств и методов интеллектуализации сверхсложных систем, взаимоувязанного применения элементов технологий автономного, распределённого, децентрализованного и группового управления, значительного увеличения динамики цифровой трансформации, относимые в рассматриваемых условиях к совокупности актуальных вопросов сбора, обработки, передачи, хранения и защиты больших массивов данных – в общих прикладных задачах цифрового моделирования и предиктивного управления, представляют собой одно из наиболее востребованных направлений и перспективных предметных областей исследования операций, многоаспектного моделирования и оценивания эффективности применения *робототехнических комплексов* (РТК) различного назначения [1–4]. *Цель исследования* – обоснование структурной модификации «предобученных на основе теории» (TTNs) / физико-информированных (PINNs) искусственных нейронных сетей на примере командно-сигнальной системы заданной структуры применения *автономных и полуавтономных функций*. Совокупность взаимоувязанных на текущем этапе программно реализуемых решений включает предварительные результаты предиктивного моделирования посредством оценивания возможностей адаптации сигнально-кодовых конструкций в аспекте изучения характерных особенностей межмашинного обмена для группового управления РТК различного назначения.

Рассматриваемая необходимость, прежде всего, находит себя в приложении приёмов многоаспектного моделирования, которые сориентированы на всесторонний учёт различных сторон и уровней системы для решения сложных задач *на ранних стадиях жизненного цикла*. В частности, речь идёт о сквозной интеграции и комплексировании данных РТК при обосновании унифицированных, более универсальных моделей для проектирования и упреждающего внедрения интеграционных компонентов в разнообразных сценариях использования таковых в сочетании с гибридными моделями технологии *анализа среды функционирования* [1–13].

Основное содержание исследований и предпроектных решений на текущем этапе включает:

исследование новых возможностей интегрированных программных решений в интересах обеспечения эффективного применения элементов передовых информационных технологий, объединяющих используемые [и иной раз принципиально разные] компоненты и платформы;

обоснование и специализация интеграционных компонентов, реализующих прогнозное моделирование в целях улучшения качества используемых комплексных моделей (прикладные аспекты цифрового моделирования и предиктивного управления), позволяющих предсказывать опорные варианты поведения системы и адаптировать их в относительно новых условиях;

оценка потенциала адаптации сигнально-кодовых структур при изучении применимости одновременно двух и более способов передачи командно-сигнальной информации и / или обмена служебными данными между объектами управления в группах с учётом характерных особенностей межмашинного обмена (*Machine-to-Machine*, M2M) / модельного задела [5–10].

Ключевые вопросы комплексного исследования, выносимые к обсуждению на секции:

- Проблемы и гипотетические комплексные решения при междисциплинарной интеграции данных [большие и сверхбольшие массивы данных]: Необходимость объединения смежных методов и подходов в области моделирования *конвергентных технологий* искусственного интеллекта, управления и связи при формировании информационно-аналитических моделей, адекватных требованиям современных систем автоматизации (автономного управления).

- Системообразующие архитектурные и технологические основы сквозной интеграции при *программно-алгоритмической реализации* обосновываемого подхода в целом: Разработка унифицированных архитектурных решений для связи и обмена данными между различными компонентами РТК, обеспечивающих масштабируемость и гибкость, а также согласование форматов данных и протоколов (включая ряд особенностей межмашинного взаимодействия).

- Аспекты цифрового моделирования вариантов опорных сценариев на этапе предпроектных исследований: Построение прогнозных моделей, позволяющих качественно и количественно оценивать поведение сложных систем, а также разрабатывать сценарии применения РТК с учётом вариативности условий эксплуатации и возможностей автоматического реагирования.

- Адаптация вариантов сигнально-кодовых структур: Исследование методов оптимизации сигнально-кодовых структур для повышения эффективности передачи данных, включая аспекты структурной скрытности и синхронизации в контексте гетерогенного управления, что может быть критично при координации и коллективном взаимодействии нескольких РТК.

- Перспективные вопросы целенаправленной специализации нейросетевых технологий: Совершенствование «*предобученных на основе теории*» (TTNs) искусственных нейронных сетей, в рамках внедрения *физико-информированных* (PINNs) моделей искусственных нейронных сетей, с возможностью дообучения в условиях реального времени и адаптации к изменяющимся условиям эксплуатации, повышает уровень автономии и устойчивости систем.

При этом, по сути, программная реализация развивает идеи и замысел ранее предложенного программно реализованного решения, которое было предназначено для обеспечения обработки больших и сверхбольших массивов данных в ряде прикладных задач авиакосмической промышленности, включая ряд комплексных вопросов дистанционного зондирования Земли, гидрометеорологии, моделирования и проблемно-ориентированного управления орбитальной группировкой и др. [5]. И, в соответствии с представленным общим назначением [6], модуль обработки больших массивов данных (*модуль* [5]) реализован как интеграционный компонент проектируемых авиационно-космических систем и комплексов, а также управления, передачи и обработки данных беспилотных (безэкипажных, автономных и др.) систем. Кроме того, модуль также обеспечивал реализацию дополнительных функций архивации и восстановления данных.

Входными данными опосредованного моделирования, в свою очередь, в рассматриваемых случаях является телеметрия и/или видеопоток с датчиков (камер), осуществляющих фото- или видеосъемку (видеопоток реального времени в произвольном из имеющихся в исходном наборе формате, – посредством сквозной системной интеграции). Выходными данными являются преобразованные изображения или видеопоток в виде бинарного потока данных специального формата, применимого исключительно для возможностей этого программного средства [5].

Модуль изначально преимущественно был ориентирован на реализацию информационных процессов *пертинентной обработки* графического контента в контуре управления авиационно-космических комплексов в интересах дальнейшего совершенствования средств и методов в основе перспективных систем для всестороннего обеспечения *Международной Хартии по космосу и крупным катастрофам* (англ. *International Charter on Space and Major Disasters*) [5].

Так, прежде всего, выделим обоснование вариантов модификации и совершенствования способа передачи сообщений *без этапности установления соединений* [6]: при исследовании новых возможностей организации информационного обмена между элементами моделируемой сети (*интерконнекта*) были выделены частные вопросы адаптации в процессе решения задачи самоорганизации связей путём изменения числа взаимодействующих элементов, а также управления приоритетом таковых элементов. Здесь, для модельных объектов управления и соответствующих критериев эффективности *неокончательных решений* [2, 6, 12], – изучается фазовый объём «*трубки решений*», развивающий далее предшествовавшие патентоспособные решения в интересах выработки предложений по формированию способов передачи данных при перераспределении задач / многомасштабных гетерогенных вычислительных ресурсов [6].

В ходе цифрового моделирования, проектирования и программно-технической реализации элементов перспективных систем цифровой обработки сигналов и изображений в рамках комплексного проекта рассматривается следующая совокупность интеграционных компонентов:

Итак, первичная модель [7] относится к области устройств, характеризующихся оптическими средствами измерения, и может быть использована для определения величины кратчайшего пути в графе. Преимущество первичной модели – расширение функциональных возможностей устройства реализацией возможности определения величины кратчайшего пути в графе [7].

Производная вторичная модель [8] относится к средствам управления потоками команд в командно-сигнальных системах сложной структуры. Значимый результат данной вторичной модели [8] заключается в определении промежуточных узлов, непосредственно связанных с управляющим центром и исполнительным звеном в командно-сигнальной системе заданной структуры. Рассматриваемая командно-сигнальная система может представлять собой сложную структуру связей между управляющим центром, промежуточным узлом и исполнительным звеном [8]; – причём, моделируемый управляющий центр формирует команды, передаваемые через промежуточные узлы исполнительному звену, при этом может существовать множество возможных путей, обеспечивающих связность моделируемых элементов такой сложной сети.

Следующая группа сопутствующих программно-технических решений [9, 10] относится к области контроля и управления распределенными многоуровневыми системами. Технический результат данного решения [9] заключается в снижении вероятности блокировки системы путём уменьшения передаваемого объёма управляющей информации за счёт её перераспределения по уровням (контурам) управления. Это решение преимущественно достигается тем, что распределенный контроль и адаптивное управление многоуровневой системой реализуют принципы построения адаптивной системы управления [9]: при нормальном функционировании распределенная структура просматривается широким оперативным полем с малой решающей способностью, достаточной для обнаружения локального нарушения режима; в дальнейшем – происходит сужение оперативного поля в окрестности нарушения, которое просматривается более детально, и, путём более детализированного анализа, производится выявление характера (распознавание), что выполняется в комплексном блоке принятия решения (реализующее данный способ устройство имеет вид блока принятия решения о состоянии многоуровневой системы и работает в трёх режимах: текущего контроля, оценки ошибок контроля и обучения).

В завершении, одно из предложенных обновлённых решений [10] относится к процедуре контроля и мониторинга и ориентировано на использование как для сетевого мониторинга технического состояния элементов *информационно-телекоммуникационных сетей* (ИТКС), так и для мониторинга состояния всей сети в целом. Сущность приводимого решения [10] заключается в сборе, накоплении и интеллектуальном анализе измерительной информации о техническом состоянии элементов сети (сетевых устройств, каналов, путей, маршрутов, узлов, сегментов, подсетей) соответственно на аппаратном, канальном и структурном уровнях ИТКС, выявлении (идентификации) их аномальных состояний и предоставлении в систему поддержки принятия решения расчётных значений достаточности задействованных ресурсов каждого уровня для устранения аварийных отказов, нагрузок и режимов работы путём регулировок параметров сетевых элементов или их замены на резервные, а также наращивании сетевых структур в случае деградации ИТКС. Способ интеллектуального мониторинга состояния ИТКС для недопущения деградации сети осуществляется с использованием нескольких этапов [10]: предварительного этапа, на котором происходит задание исходных данных мониторинга; первого этапа, на котором формируют контур наблюдения и осуществляют сбор измерительной информации с сетевых устройств, сетевых соединений и сетевых структур распределенной ИТКС, а также передачу её от серверов мониторинга сети на серверы ситуационного центра; второго этапа, на котором оценивают техническое состояние элементов сети в серверах ситуационного центра по оценке надёжности, оценке помехоустойчивости и оценке живучести (связности) сети; третьего этапа, в ходе которого формируется контур управления сетевыми элементами путём перераспределения локальных параметрических (регулировка параметров), телекоммуникационных (корректировка уровней передачи в каналах связи) и структурных (перемаршрутизация) ресурсов сети, а также дополнительных аппаратных (резервирование сетевых устройств), канальных (резервирование каналов связи) и топологических (изменение топологии) ресурсов сети; завершающего этапа, заключающегося в актуализации базы состояния

системы мониторинга, прогнозирования состояния элементов сети, формировании различных видов отчетов (протоколов) и формировании «*витрины данных*» по всему циклу мониторинга. Таким образом, ключевой технический результат здесь – интеллектуализация процедуры сетевого мониторинга на основе мультиагентного подхода [9, 10]. При этом, так называемая «*витрина данных*» представляет собой немаловажную составляющую в практике прикладной реализации рассматриваемых интеграционных компонентов в составе проблемно-ориентированной платформы («*интеллектуального стенда*»). И, далее, путём обоснованного применения приводимых в докладе приёмов сквозной системной интеграции и дальнейшей адаптации пула достигаемых частных результатов предполагается создание *прототипа* универсальной платформы цифрового моделирования, которая позволит обеспечить сквозное выполнение процессов сбора, обработки, анализа и использования данных РТК. Это, в свою очередь, обеспечит повышение эффективности, надёжности и гибкости сложных систем, что является критически важным для их реализации в реальных условиях эксплуатации [11–13].

Преимущество подхода – реализация новых возможностей «*высокоуровневого*» итеративного исследования сценариев группового применения РТК в интересах детализированной осведомленности лиц, принимающих решение, в ходе организации и проведения предпроектных исследований посредством сквозной интеграции больших массивов данных для структурно-функциональной коррекции обучения TTNs / PINNs на примере командно-сигнальной системы заданной структуры применения автономных и полуавтономных функций.

В числе наиболее интересных направлений специализации при разработке и исследовании существенных свойств интеграционных компонентов представляется возможным выделить соответствующие вопросы *дифференциальной приватности* на базе теоретических положений с применением элементов научно-методического аппарата теории графов в сочетании с идеями формализации в терминах *перколяции*, в том числе в приложении к формированию адаптивных сигнально-кодовых конструкций для рассматриваемых новых задач комплексной проблематики многоаспектного цифрового моделирования новых *автономных и полуавтономных функций*. В свою очередь, *симуляционные среды* также представляют особый интерес в свете возможностей популяризации передовых достижений *автономных робототехнических комплексов*, всесторонне рассмотрение которых – вкупе со многими смежными вопросами (*прим*.: см. также https://www.ipu.ru/robot). Данный подход также во многом способствует развитию теоретической базы и перспективных способов, предназначенных для решения задач междисциплинарной интеграции в рамках парадигмы формирования *цифровых двойников,* глубокой модернизации комплексных систем обмена сигналами с применением M2M и адаптивного машинного обучения, что также является одним из существенных преимуществ и предпосылкой для новых рекомендаций по *стандартизации* в робототехнике и автоматизации.

Литература

1. Рожнов, А. В. Исследование гибридных моделей анализа среды функционирования при описании восприятия действий просьюмеров с использованием микродиректив // XIV Всероссийское совещание по проблемам управления : сборник научных трудов, Москва, 17–20 июня 2024 года. – М.: Институт проблем управления им. В.А. Трапезникова РАН, 2024. – С. 1239-1242. – EDN DFDBDR.
2. Кузнецова, В. Л. Самоорганизация в технических системах / В. Л. Кузнецова, М. А. Раков; АН УССР, Физ.-мех. ин-т им. Г. В. Карпенко. - Киев: Наук. думка, 1987. - 196 с.
3. Рожнов, А. В. Особенности зарубежных представлений и технической политики по разработке и использованию автономных и полуавтономных функций сложных технических систем (по материалам открытой печати) // Перспективные системы и задачи управления : Материалы XVIII Всероссийской научно-практической конференции и XIV молодежной школы-семинара, 03–07 апреля 2023 года. – Таганрог, 2023. – С. 118-129. – EDN DRWSKA.
4. Рожнов, А. В. К вопросу о сложности и моделировании задач командования и управления // Перспективные системы и задачи управления : Материалы XIX Всероссийской научно-практической конференции и XV молодежной школы-семинара, 01 апреля 2024 года. – Таганрог, 2024. – С. 23-28. – EDN PTXMGD.

5. Свидетельство о государственной регистрации программы для ЭВМ № 2018618078. Модуль обработки больших массивов данных : № 2018615273 : заявл. 24.05.2018 : опубл. 09.07.2018 / В. И. Гончаренко, А. А. Моисеев, В. В. Цыпелев [и др.] ; заявитель ФГБОУ ВО «Московский авиационный институт (национальный исследовательский университет)». – EDN CKXQZP.
6. Рожнов, А. В. Исследование новых возможностей совершенствования способа передачи сообщений без этапности установления соединений // Устойчивость и колебания нелинейных систем управления (конференция Пятницкого) : Материалы XVI Международной научной конференции, Москва, 01–03 июня 2022 года / Под общей редакцией В.Н. Тхай. – М.: Институт проблем управления им. В.А. Трапезникова РАН, 2022. – С. 378-380. – EDN QWDJVK.
7. Патент на полезную модель № 61409 U1, МПК G01B 11/02. Устройство для определения величины кратчайшего пути в графе : № 2006139635/22 : заявл. 09.11.2006 : опубл. 27.02.2007. – EDN ZPHMDZ.
8. Патент на полезную модель № 231144 U1, МПК H04L 45/00. Устройство определения промежуточных узлов, непосредственно связанных с управляющим центром и исполнительным звеном в командно-сигнальной системе заданной структуры : заявл. 03.05.2024 : опубл. 13.01.2025. – EDN TPNCCR.
9. Патент № 2450335 C1, МПК G06F 15/00, G05B 23/02. Способ распределенного контроля и адаптивного управления многоуровневой системой и устройство для его осуществления : № 2011128648/08 : заявл. 11.07.2011 : опубл. 10.05.2012. – EDN ZGIWCD.
10. Патент № 2841123 C1, МПК G05B 19/048, H04L 41/06. Способ интеллектуального мониторинга состояния информационно-телекоммуникационной сети : заявл. 12.11.2024 : опубл. 03.06.2025. – EDN SPJHJD.
11. Саломатин, А. А. Влияние метеорологических характеристик на полет беспилотного летательного аппарата // XIV Всероссийское совещание по проблемам управления : сборник научных трудов, Москва, 17–20 июня 2024 года. – М.: Институт проблем управления им. В.А. Трапезникова РАН, 2024. – С. 1190-1193. – EDN LCEOMD.
12. Рожнов, А. В. Обоснование интеграционных решений при комплексировании сервисной информации в задачах интеллектуальной обработки и анализа гетерогенных данных среды функционирования киберфизических систем / А. В. Рожнов, О. М. Гергет // Решетневские чтения : Материалы XXVIII Международной научно-практической конференции, посвященной 100-летию со Дня рождения генерального конструктора ракетно-космических систем академика Михаила Федоровича Решетнева: 2-х частях. – Красноярск: СибГУ, 2024. – С. 437-440. – EDN MLXZTU.
13. Саломатин, А. А. Анализ структур систем информационной поддержки транспортировки грузов с помощью гетерогенной группы БПЛА с учётом метеорологических условий // Вестник ДонНУ. Серия Г: Технические науки. – 2024. – № 3. – С. 92-101. – DOI 10.5281/zenodo.14018632. – EDN BOJJIQ.

INVESTIGATION OF ROBOTICS COMPLEX APPLICATION SCENARIOS IN THE DIGITAL MODELING OF SEMI-AUTONOMOUS FUNCTIONS

Salomatin A.A., Rozhnov A.V., Zhuravleva N.G.

ICS RAS

Multiaspect modeling in the context of digital transformation imposes new requirements for the justification and implementation of interdisciplinary results in solving complex problems of organizing and conducting pre-design studies of scenarios for the use of robotic complexes (RTCs) for various purposes. The aim of the work is to study problematic issues of end-to-end integration of RTC data in digital modeling, design and implementation of integration components in the context of semi-autonomous functions. The set of interconnected software-implemented solutions at the current stage includes a number of applied results of predictive modeling, evaluation of potential for adapting signal-code structures, including the characteristic features of machine-to-machine exchange for group control, as well as modifications of theory-trained neural networks.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ СПЕКТРАЛЬНОГО РАЗЛОЖЕНИЯ МАТРИЦЫ КОРРЕЛЯЦИИ ДЛЯ РЕАЛИЗАЦИИ НА ПЛИС

инж. Стародубов Д.А., к.т.н. Глухов А.Ю., инж. Кириллова К.В., к.т.н. Петров С.А.

ФГУП Всероссийский научный исследовательский институту автоматики им. Духова

В задачах пеленгования, локации, мониторинга окружающей обстановки выделяется отдельный класс методов сверхразрешения, в основе которых лежит выделение подпространств сигнала и шума с использованием разложения матрицы корреляции на собственные числа и вектора. Ввиду широкого распространения автономных малогабаритных мобильных систем, задача исследования аппаратной реализации методов является весьма актуальной. В работе проведен обзор методов нахождения собственных чисел и векторов матрицы корреляции, а также их сравнение в части оценки вычислительной сложности и точности. Получены графики зависимостей количества операций и максимального значения нормы невязки в выбранных методах от ОСШ. Сделан вывод о перспективах применения методов для реализации на ПЛИС.

Введение

Задачи вычисления собственных чисел и собственных векторов матриц занимают важное место в современной теории обработки сигналов, радиотехнике и вычислительной математике. Особую актуальность данные задачи приобретают в многоканальных системах приема и обработки сигналов, где анализ спектральных свойств матриц корреляции (МК) и ковариации является основой для выделения информативных подпространств и последующей оценкой параметров сигналов. К таким системам, в частности, относятся системы с приемными антенными решётками (АР) с ультракороткой базой (УКБ), широко применяемые в локации, навигации, мониторинге, где важно точно определять направление прихода сигнала.

При применении АР с УКБ различие фазовых задержек между элементами существенно уменьшается, что приводит к ухудшению углового разрешения при использовании классических методов пеленгации. В этих условиях особое значение приобретают методы сверхразрешения, основанные на собственном разложении матрицы корреляции принимаемых сигналов, например MUSIC. Такие алгоритмы используют разделение пространства наблюдений на сигнальное и шумовое подпространства, формируемые на основе собственных векторов матрицы, что напрямую связывает точность и устойчивость оценки направления прихода сигнала с качеством и быстродействием вычисления собственных чисел и собственных векторов.

Многоканальные системы цифровой обработки сигналов (ЦОС), реализованные на базе антенных решеток, предъявляют повышенные требования к вычислительной производительности и энергоэффективности. Традиционные аппаратные решения, основанные на цифровых сигнальных процессорах (ЦСП) имеют существенные ограничения в условиях многоканальной обработки из-за последовательного выполнения арифметических операций. Попытки распараллеливания вычислительных задач на многоядерных ЦСП сопровождаются значительным увеличением энергопотребления, что критично для автономных малогабаритных систем. В то же время в алгоритмах сверхразрешения можно выделить параллельную структуру, позволяющую распределить вычислительную нагрузку между независимыми операционными блоками. Это делает целесообразным использование программируемых логических интегральных схем (ПЛИС) [1-4].

Целью работы является выбор метода вычисления собственных чисел и векторов матрицы корреляции для применения в системах реального времени на базе ПЛИС.

Постановка задачи

Рассмотрим задачу определения направления прихода узкополосного сигнала с частотой f на АР, состоящую из M приемных элементов. Пусть каждый приемный элемент можно считать материальной точкой, а фронт прихода волны, приходящий на каждый преобразователь, плоским (дальняя зона).

Если радиус-вектора положений преобразователей АР $r_m \in \mathbb{R}^3, m = 1, \dots, M$, то направляющий вектор АР для волны, пришедшей с горизонтального угла θ и вертикального угла φ

$$v(\theta,\varphi) = \left[e^{-jk^T r_1}, e^{-jk^T r_2}, \dots, e^{-jk^T r_M}\right]^T$$

где $k = -\frac{2\pi}{\lambda}\begin{bmatrix}\sin\theta\cdot\cos\varphi\\ \sin\theta\cdot\sin\varphi\\ \cos\theta\end{bmatrix}$ – волновое число, соответствующее частоте f, λ – длина волны.

Если АР принимает сигналы от D источников, то формируется направляющая матрица

$$\vartheta = [v(\theta_1,\varphi_1), v(\theta_2,\varphi_2), \dots, v(\theta_D,\varphi_D)]^T$$

Пусть $s_d[k]$ представляет собой дискретный сигнал, содержащий K отсчетов $k = 1..K$ от одного из D источников, $x_d[k]$ – дискретные отсчеты сигнала, принимаемого одним из приемных элементов АР, $n_d[k]$ – дискретные отсчеты шумовой составляющей на каждом из M приемных элементов. Объединим дискретные отсчеты в вектора s_d, x_d, n_d с размерностями $[1 \times K]$. Далее вектора s_d от всех источников – в матрицу S с размером $[M \times K]$ и векторы x_d, n_d всех приемных элементов АР – в матрицы сигналов X и N с размером $[M \times K]$ соответственно. Считаем, что шумовая составляющая образуется вследствие воздействия аддитивного белого Гауссова шума (АБГШ). Тогда математическая модель сигналов, принятых каждым из M приемных элементов АР

$$X = \vartheta \cdot S + N.$$

Ключевую роль в методах разделения пространства наблюдений на сигнальное и шумовое подпространства играет матрица корреляции в частотной области [5]. Для ее получения необходимо перевести исходные сигналы в частотную область с помощью преобразования Фурье $X \to \hat{X}$ и осуществить матричное перемножение

$$R_{xx} = \hat{X} \cdot \hat{X}^H. \quad (1)$$

Матрица корреляции R_{xx} является эрмитовой и положительно определенной, что позволяет выполнить ее собственное разложение

$$R_{xx} = V \Lambda V^H,$$

где $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_M)$ – диагональная матрица собственных чисел, упорядоченных по убыванию, $V = [V_s\ V_n]$ – матрица собственных векторов, образующих сигнальное V_s и шумовое V_n подпространства.

Дальнейший обзор методов вычисления собственных чисел и векторов эрмитовой матрицы А вида (1) выполнен в контексте рассмотренной задачи.

Обзор методов вычисления собственных чисел и векторов

В литературе [6,7 и др.] для нахождения собственных чисел и векторов описываются прямые и итерационные методы. Прямые методы основаны на вычислении коэффициентов характеристического многочлена, корни которого являются собственными числами. Далее, в зависимости от метода (Леверье-Фаддеева, Крылова и др.) производится расчет соответствующих собственных векторов. Плохая численная обусловленность [8], а также реализация в условиях ограниченной разрядности чисел приводят к невозможности применения данных методов для решения практических задач на ПЛИС. Поэтому интерес представляет рассмотрение итерационных методов.

Среди них выделяются Якоби и QR, разложение в котором может быть реализовано такими методами как модифицированный метод ортогонализации Грама-Шмидта, вращения Гивенса, отражения Хаусхолдера. В данной работе QR-разложение осуществляется модифицированным методом Грама-Шмидта, который согласно [9] численно идентичен вращениям Гивенса, обеспечивая при этом тот же уровень численной стабильности и точности.

В свою очередь, методы Якоби и QR основаны на применении последовательных унитарных преобразований, что позволяет привести матрицу к диагональному или треугольному виду. Диагональные элементы такой матрицы являются собственными числами исходной матрицы.

Метод Якоби основан на последовательной диагонализации эрмитовой матрицы путем последовательного зануления ее внедиагональных элементов. Для исходной матрицы A строится последовательность эрмитовых матриц $A^{(k)}$

$$A^{(k+1)} = U_k^H A^{(k)} U_k,$$

где k – номер итерации метода Якоби, U_k – комплексная унитарная матрица простого поворота [10], на нулевой итерации принимается $A^{(0)} = A$.

Матрица U_k имеет вид

$$U_k = \begin{bmatrix} I & & & & \\ & cos\varphi^{(k)} & & -e^{i\psi^{(k)}} \sin\varphi^{(k)} & \\ & & I & & \\ & e^{-i\psi^{(k)}} \sin\varphi^{(k)} & & cos\varphi^{(k)} & \\ & & & & I \end{bmatrix} \begin{matrix} \\ i \\ \\ j \\ \\ \end{matrix}$$

где углы

$$\varphi^{(k)} = \frac{1}{2} arctg\left(\frac{2a_{ij}^{(k)}}{a_{ii}^{(k)} - a_{jj}^{(k)}}\right), \psi^{(k)} = \arg\left(a_{ij}^{(k)}\right),$$

а $e^{i\psi^{(k)}}$ и $e^{-i\psi^{(k)}}$ выражается из показательной формы записи комплексного числа

$$e^{i\psi^{(k)}} = \frac{a_{ij}^{(k)}}{\left|a_{ij}^{(k)}\right|}, e^{-i\psi^{(k)}} = \frac{\left(a_{ij}^{(k)}\right)^*}{\left|a_{ij}^{(k)}\right|}.$$

Индексы i, j определяются таким образом, чтобы модуль внедиагонального элемента $\left|a_{ij}^{(k)}\right|$ был наибольшим среди всех элементов матрицы $A^{(k)}$ при $i < j$.

Матрица собственных векторов определяется как произведение всех унитарных матриц поворота. На каждой итерации обновление матрицы собственных векторов осуществляется согласно $V_k = V_{k-1}U_k$, где на начальной итерации $V_0 = U_0$.

В результате последовательных унитарных преобразований норма внедиагональных элементов матрицы $A^{(k)}$ уменьшается, и матрица стремится к диагональному виду.

Критериями остановки вычислений являются условия:

– максимальный по модулю внедиагональный элемент меньше заданной точности ε;

– достижение алгоритмом максимального заданного количества итераций.

В QR-алгоритме приведение исходной матрицы A к верхнетреугольному виду осуществляется за счет последовательных QR-разложений матрицы A

$$A^{(k)} = Q^{(k)}R^{(k)}$$

и последующем обновлении матрицы $A^{(k+1)}$ по формуле:

$$A^{(k+1)} = R^{(k)}Q^{(k)},$$

где $Q^{(k)}$ – унитарная матрица, $R^{(k)}$ – верхняя треугольная, на нулевой итерации принимается $A^{(0)} = A$.

Для QR-разложения используется модифицированный метод ортогонализации Грама-Шмидта. Он заключается в последовательной ортогонализации столбцов матрицы с вычитанием проекций на уже построенные ортонормированные векторы. На каждом шаге вычисляется норма текущего столбца $a_j, j = 1, \dots, M$ и формируется ортонормированный вектор q_j путем нормировки a_j. Ортогонализация последующих столбцов $a_i, i = j + 1, \dots, M$ относительно q_j осуществляется путем последовательного вычитания проекций.

Таким образом, для столбцов $a_1, \dots, a_M$ матрицы $A^{(k)}$ вычисляется ортонормированный базис $q_1, \dots, q_M$, образующий столбцы матрицы $Q^{(k)}$, при этом $q_j = \frac{a_j}{\|a_j\|}, r_{jj} = \|a_j\|$, а элементы матрицы $R^{(k)}$ и обновление столбцов матрицы $A^{(k)}$ определяются по формулам

$$r_{ji} = q_j^H a_i \text{ при } i > j, \; a_i = a_i - r_{ji}q_j.$$

В результате итерационного процесса матрица $A^{(k)}$ сходится к верхнетреугольному виду. Матрица собственных векторов определяется как произведение всех унитарных матриц. На каждой итерации обновление матрицы собственных векторов осуществляется согласно $V^{(k)} = V^{(k-1)}Q^{(k)}$, где на начальной итерации принимается $V^{(0)} = Q^{(0)}$.

Критерии остановки вычислений аналогичны критериям в методе Якоби. Однако, так как матрица $A^{(k)}$ сходится к верхнетреугольному виду, то в первом критерии проверяется максимальный по модулю внедиагональной элемент ниже главной диагонали.

Таким образом, оба рассмотренных метода представляют интерес. Для оценки возможности их реализации в ПЛИС необходимо оценить точность и вычислительную сложность в контексте задачи определения собственных чисел и векторов в условиях работы с аддитивным белым гауссовым шумом (АБГШ).

Выбор метода для реализации на ПЛИС

Так как ПЛИС, ввиду их параллельной архитектуры, целесообразно применять в системах реального времени, то для сравнения методов Якоби и QR в качестве метрики оценки вычислительной сложности было выбрано количество операций. Результаты оценки представлены на рисунке 1. Ввиду наличия АБГШ и итерационного характера методов с критерием выхода по заданному ε производилось усреднение количества операций по 100 реализациям для соответствующего значения ОСШ.

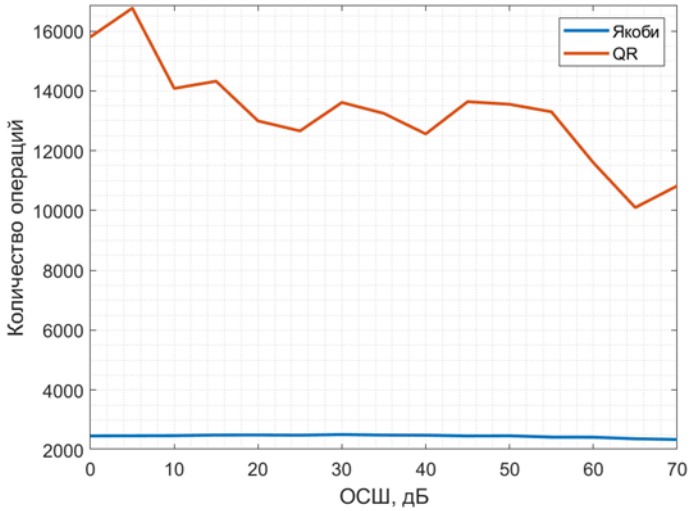
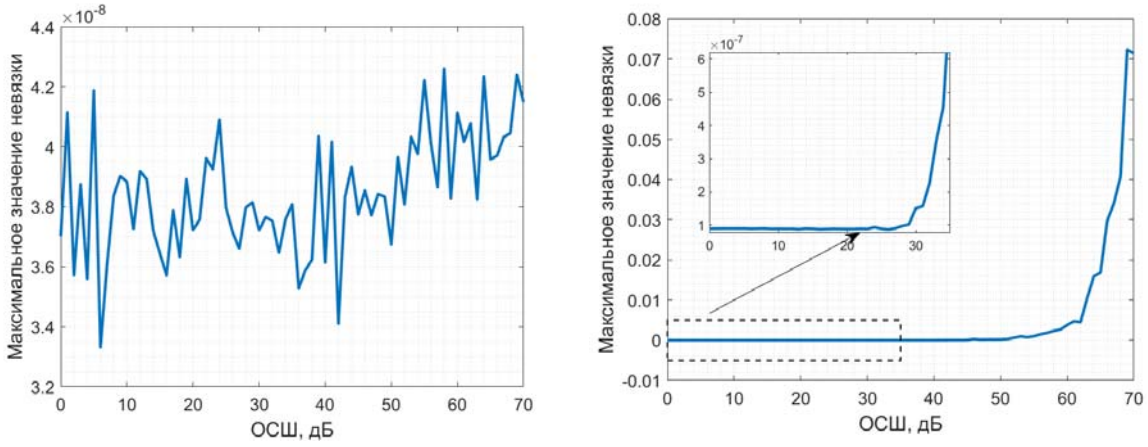


Рисунок 1. График зависимости количества операций в алгоритмах Якоби и QR от ОСШ

Видно, что количество операций в методе Якоби, в отличие от QR, слабо зависит от ОСШ. Также метод Якоби требует примерно в 6-8 раз меньше операций, чем QR, что делает его более подходящим для реализации в ПЛИС.

В качестве метрики оценки точности для каждой из 100 реализаций ОСШ было выбрано максимальное значение нормы невязки $\max_i \|Ax_i - \lambda_i x_i\|$, где $i = 1, \dots, M$. Далее производилось усреднение метрики по 100 реализациям для соответствующего значения ОСШ. Результаты оценки представлены на рисунке 2.



а) б)

Рисунок 2. Графики зависимости максимального значения нормы невязки в алгоритмах Якоби (а) и QR (б) от ОСШ

Видно, что для метода Якоби значение нормы невязки не превышает $4{,}4 \cdot 10^{-8}$ во всем задаваемом диапазоне входного ОСШ. В свою очередь, для метода QR с увеличением ОСШ это значение увеличивается, приближаясь к значению $8 \cdot 10^{-2}$. В диапазоне значений ОСШ от 0 до 35 норма невязки не превышает значения $6 \cdot 10^{-7}$, которое больше аналогичного значения, полученного при реализации метода Якоби.

С учетом работы в реальном времени меньшее количество затрачиваемых операций в методе Якоби, а также меньшее значение ошибки в определении собственных чисел и векторов, делает метод Якоби более подходящим для реализации на ПЛИС.

Заключение

В работе рассмотрены методы Якоби и QR для нахождения собственных чисел и векторов, а также проведено их сравнение в части оценки вычислительной сложности и точности, на основе чего сделан выбор метода Якоби для аппаратной реализации на ПЛИС.

Полученные результаты могут быть полезны при разработке алгоритмов сверхразрешения типа MUSIC в условиях антенных решеток с ультракороткой базой на базе ПЛИС, работающих в масштабе реального времени для решения задач локации, навигации, мониторинга окружающего пространства.

Литература

1. Chen K. A low complexity, high throughput DOA estimation chip design for adaptive beamforming // Electronics. – 2020. – Т. 9. – №4. – С. 641.

2. Hussain A. FPGA-based hardware implementation of computationally efficient multi-source DOA estimation algorithms // IEEE access. – 2017.

3. Assoa A.S. A scalable platform for single-snapshot direction of arrival (DOA) estimation in massive MIMO systems // Lakes Symposium on VLSI 2023. – 2023. – С. 631-637.

4. Chen H. An efficient hardware accelerator for the MUSIC Algorithm // Electronics. – 2019. –Т. 8. – №.5. – С. 511.

5. Петров С.А., Глухов А.Ю., Стародубов Д.А. Сложнофункциональный блок для вычисления матрицы корреляции в частотной области // Цифровая обработка сигналов. 2025. №2. С. 15-20.

6. Березин И.С., Жидков Н.П., Методы Вычислений. Т.2. – Москва : Физматлит, 1959. – 620 с.

7. Голуб Д., Лоун В., Матричные вычисления, Пер. с англ. – Москва : Мир, 1999. – 548 с.

8. Trefethen L.N., Bau D. Numerical Linear algebra. – Philadelphia : Society for Industrial and Applied Mathematics, 1997. – 361 p.

9. Ganchosov P.N. et al. FPGA Implementation of Modified Gram-Schmidt QR-Decomposition // Proceedings of the 3rd HiPEAC workshop on reconfigurable computing. – 2009. – pp.41-51.

10. Hari V., Kovac E.B. On the convergence of complex Jacobi methods // Linear and multilinear algebra. – 2019.

COMPARATIVE ANALYSIS OF THE SPECTRAL DECOMPOSITION OF THE CORRELATION MATRIX FOR FPGA IMPLEMENTATION

Starodubov D.A., Glukhov A.Yu., Kirillova K.V., Petrov S.A.

FSUE VNIIA

In the fields of direction finding, location, and environmental monitoring, a distinct class of super-resolution methods is distinguished. These methods are based on the extraction of signal and noise subspaces using the decomposition of the correlation matrix into eigenvalues and eigenvectors. Due to the widespread use of autonomous small-sized mobile systems, the task of studying the hardware implementation of these methods is highly relevant. This paper provides a review of methods for finding the eigenvalues and eigenvectors of the correlation matrix, as well as a comparison of their computational complexity and accuracy. Graphs of the number of operations and the maximum value of the residual norm in the selected methods versus the SNR are obtained. A conclusion is drawn regarding the prospects for applying these methods to the FPGA.

АЛГОРИТМЫ ПЛАНИРОВЩИКА СИНХРОННЫХ ПОТОКОВ В ГРАФИЧЕСКИХ ПРОЦЕССОРАХ

асп. Титаев Д.В.[1], доц. Бочаров Ю.И.[1], асс. Звягин А.А.[1], асп. Ковшаров И.Д.[1], доц. Бутузов В.А.[1]

[1] Национальный исследовательский ядерный университет «МИФИ»

В работе представлены результаты анализа современных аппаратных методов планирования групп синхронных потоков (варпов) и их применимости в графических процессорах (ГП). Предложены классификация и сравнительное сопоставление актуальных алгоритмов планирования варпов. Для каждого подхода кратко изложены принцип работы, ключевые достоинства и ограничения, а также типичные сценарии применения. Показано, что эффективность алгоритмов определяется как характером нагрузки, так и особенностями подсистемы памяти.

Введение

Графические процессоры (ГП) – это вычислительные устройства с массовым параллелизмом, предназначенные для одновременной обработки тысяч элементов данных. Для эффективного выполнения вычислений современные ГП используют архитектуру SIMT (Single instruction multiple threads – одна инструкция, множественные потоки). Устройство выполняет одну и ту же последовательность инструкций большим числом потоков [1]. Группу потоков с единой последовательностью инструкций называют волновым фронтом («wavefront», в терминологии AMD) или варпом («warp», в терминологии NVIDIA). Каждый варп состоит из фиксированного числа потоков, которые разделяют один программный счётчик и исполняются синхронно в одном вычислительном блоке – потоковом мультипроцессоре (ПМП), который в англоязычной литературе называют Streaming Multiprocessor в терминологии компании NVIDIA или Compute Unit в терминологии компании AMD.

Несмотря на тысячи параллельных потоков, производительность ГП может оказаться существенно ниже ожидаемой из-за невозможности отправить очередной варп на исполнение. Ключевыми факторами являются задержки в работе исполнительных блоков, а также задержки при поступлении данных из памяти и от других функциональных блоков [2]. Планировщик варпов – это аппаратный блок в каждом потоковом процессоре, определяющий очередность выполнения множества активных варпов. В соответствии с алгоритмом планировщика в каждом такте выбирается варп для выдачи следующей инструкции. Планирование очередности варпов необходимо как для минимизации блокировок конвейера, так и для эффективного использования вычислительных ресурсов.

Цель настоящей работы заключается в систематизации алгоритмов аппаратного планирования варпов и определении критериев их выбора в зависимости от характера нагрузки.

2. Модель потокового мультипроцессора и сигналы для планировщика варпов

Архитектура ГП существенно отличается от архитектуры других вычислительных устройств, например микроконтроллеров. Графический процессор состоит из множества вычислительных кластеров, каждый из которых имеет доступ к кэш-памяти второго уровня (L2-кэш) и к глобальной памяти, а каждый кластер включает несколько ПМП, как показано на рисунке 1 [3, 4].

Потоковый мультипроцессор состоит из нескольких блоков: кэша инструкций, планировщика варпов, большого регистрового файла, блоков чтения и записи, а также множества ядер [5-8]. Ядро включает в себя блоки работы с целыми числами и с числами с плавающей точкой. В модели SIMT потоки объединяются в варпы, и для всех потоков одного варпа используется единый программный счётчик; для планировщика варп является минимальной единицей планирования.

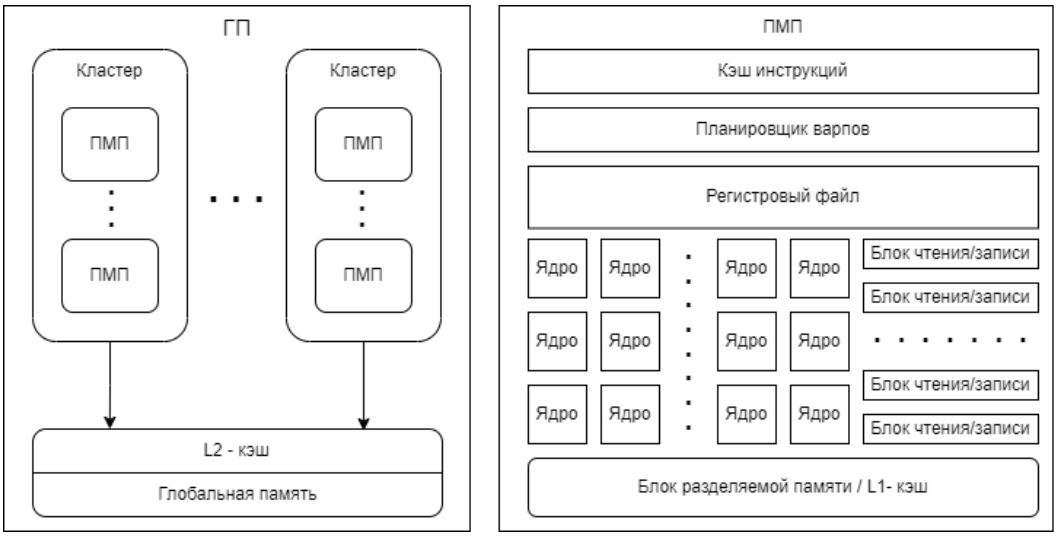


Рисунок 1 – Архитектура ГП и структура ПМП

3. Функции планировщика

Работа планировщика реализуется в виде повторяющегося цикла: на каждом такте из множества резидентных варпов выбирается варп, готовый к выдаче следующей инструкции. Если текущий варп блокируется вследствие задержки, планировщик осуществляет переключение на другой варп с готовыми к выдаче инструкциями, стремясь скрыть задержки за счёт чередования варпов.

Понятие готовности варпа определяется не только наличием инструкции по текущему значению программного счётчика, но и возможностью её немедленного выполнения с учётом зависимостей по данным, доступности функциональных блоков, состояния кэш-подсистемы и подсистемы памяти.

4. Классы по цели оптимизации

Представленная ниже классификация объединяет алгоритмы планирования варпов по основной цели оптимизации и типу ограничений, которые они позволяют ослабить.

4.1 Справедливость и простота.

Данный класс ориентирован на обеспечение равномерного распределения вычислительных ресурсов между варпами при минимальной аппаратной и логической сложности планировщика. К числу типичных представителей относятся циклические алгоритмы SRR и LRR [9].

4.2 Локальность и пропускная способность.

Алгоритмы данного класса направлены на повышение пространственной и временной локальности, а также на увеличение пропускной способности за счёт сокращения частоты переключений между варпами и/или упорядочения активного множества варпов. К типичным представителям относятся алгоритм GTO [3], в котором при переключении выбирается дольше всего ожидающий варп, и TLS [3], основанный на разделении варпов на группы для повышения эффективности выполнения.

4.3 Снижение нагрузки на L1-кэш и уменьшение числа конфликтов.

Основной задачей данного класса является снижение числа конфликтных промахов и сохранение локальности обращений к данным при избыточном количестве активных варпов. К числу представителей относятся методы статического ограничения числа активных варпов (SWL) [4, 10], а также адаптивные подходы (CCWS, DAWS) [10], использующие анализ кэш-промахов и оценку или прогноз загрузки кэша.

4.4 Предотвращение блокировок в подсистеме памяти.

Алгоритмы данного класса направлены на то, чтобы не допускать исчерпания внутренних ресурсов, обслуживающих кэш-промахи и операции загрузки, поскольку их нехватка может препятствовать дальнейшему выполнению программы. Типичным представителем является алгоритм OAWS [11], ориентированный на контроль использования ресурсов, связанных с обработкой операций загрузки.

4.5 Синхронизация и активное ожидание.

Цель алгоритмов данного класса — сократить непроизводительные затраты, возникающие при активном ожидании на примитивах синхронизации. В качестве представителя можно выделить BOWS [12], использующий механизмы выявления варпов, не продвигающихся по выполнению, и их временного исключения из планирования.

5. Типичные алгоритмы планирования

Далее приведены наиболее показательные алгоритмы планирования варпов, иллюстрирующие практическую реализацию рассмотренных классов подходов.

5.1 Циклические алгоритмы

SRR (Static Round Robin) осуществляет обход варпов в фиксированном цикле без учёта готовности. Достоинством является минимальная аппаратная сложность и предсказуемое распределение квантов времени. Недостатком выступает низкая эффективность при наличии блокировок: готовые варпы могут простаивать, пока планировщик проходит по неготовым.

LRR (Loose Round Robin) сохраняет циклический обход, но пропускает неготовые варпы, выбирая следующий готовый варп. Это повышает утилизацию вычислительных устройств при близкой аппаратной сложности. Потенциальным ограничением является неравномерность задержек: часть варпов может получать выдачу чаще, чем другие, в зависимости от характера блокировок.

В качестве практического базового решения более предпочтителен LRR, тогда как SRR полезен главным образом как опорный эталон [9].

5.2 Жадные алгоритмы и локальность

GTO (Greedy Then Oldest) реализует жадное правило: после выбора варпа планировщик стремится выдавать ему инструкции последовательно до момента блокировки либо завершения. При необходимости смены варпа следующий кандидат выбирается по критерию максимального времени ожидания (наиболее “старый” варп). Такой механизм снижает частоту переключений, что, как правило, благоприятно сказывается на локальности и может повышать общую пропускную способность, особенно при однородных потоках исполнения.

Ключевым ограничением GTO является ухудшение справедливости: при определённых условиях задержки между варпами выравниваются хуже, поскольку один варп может слишком долго занимать очередь выдачи.

TLS (Two-Level Warp Scheduling) — это двухуровневая политика планирования варпов, в которой все активные варпы разделяются на группы, а выбор осуществляется в два шага: сначала выбирается группа, затем внутри выбранной группы применяется базовая стратегия выдачи (обычно жадная или локально-ориентированная) [3]. Такой подход уменьшает накладные расходы частых переключений между разнородными варпами и повышает устойчивость исполнения за счёт более контролируемого чередования, что в ряде случаев улучшает локальность и снижает простои.

5.3 Контроль использования кэша первого уровня

При избыточном числе активных варпов рабочие наборы данных начинают интенсивно вытеснять друг друга из L1, что приводит к росту конфликтных промахов и увеличению доли времени, затрачиваемого на ожидание данных.

SWL (Static Wavefront Limiting) ограничивает число одновременно активных варпов, даже если аппаратные ресурсы допускают большее число активных варпов. Параметр ограничения выбирается профилированием или экспериментальной настройкой [4, 10]. Достоинство SWL — минимальная аппаратная стоимость; недостаток — слабая переносимость между различными ядрами и фазами исполнения, а также риск недоиспользования ПМП при неверном выборе лимита.

CCWS (Cache-Conscious Wavefront Scheduling) использует адаптивную регулировку активности варпов на основе событий кэша [10]. Вводится аппаратный механизм, фиксирующий вытеснения и повторные обращения, что позволяет идентифицировать варпы с высокой долей конфликтных промахов. Такие варпы получают приоритет, тогда как часть варпов временно ограничивается, что снижает конкуренцию и повышает пространственную локальность. Ограничением является повышенная аппаратная сложность.

DAWS (Divergence-Aware Warp Scheduling) ориентирован на случаи выраженной нерегулярности и дивергенции обращений. В отличие от CCWS, имеющего реактивный характер, DAWS стремится прогнозировать величину дивергенции обращений: на этапе

компиляции программы анализируются загрузки и потенциальное число затрагиваемых кэш-строк, а на этапе исполнения аппаратный предсказатель оценивает, приведёт ли добавление нового варпа к переполнению эффективной кэш-ёмкости [13]. При ожидаемом ухудшении локальности выдача варпа откладывается до снижения нагрузки на кэш.

5.4 Конфликт ресурсов подсистемы памяти

Подсистема памяти ПМП ограничена внутренними структурами, обеспечивающими обработку промахов кэша; типичным примером таких структур являются MSHR (Miss Status Holding Registers), предназначенные для отслеживания незавершённых кэш-промахов и операций загрузки. При низкой пространственной локальности варпы способны генерировать большое количество параллельных обращений к памяти, что приводит к быстрому исчерпанию доступных записей MSHR. В результате новые запросы не могут быть обслужены даже для готовых варпов, что приводит к простоям и снижает общую эффективность выполнения.

Алгоритм OAWS (Memory Occlusion Aware Warp Scheduling) направлен на предотвращение исчерпания записей MSHR при выполнении загрузок с низкой пространственной локальностью [11]. Когда все записи MSHR заняты, даже готовые к выполнению инструкции загрузки не могут получить доступ к кэш-подсистеме и вынуждены ожидать освобождения хотя бы одной записи. В отличие от традиционных подходов, OAWS не допускает полного исчерпания свободных записей MSHR, поскольку динамически отслеживает степень их использования и оценивает риск возникновения блокировки по памяти. Для этого алгоритм прогнозирует, какое количество записей MSHR потребуется следующей инструкции загрузки активного варпа. Если ожидается превышение допустимого порога, выполнение данного варпа временно откладывается. Такой подход особенно эффективен для приложений с интенсивным и неоднородным характером обращений к памяти.

5.5 Активное ожидание на синхронизационных примитивах

Существенной проблемой при выполнении программ на графических процессорах является активное ожидание варпов, возникающее при использовании механизмов синхронизации, таких как атомарные операции, барьеры и иные средства координации потоков. В подобных ситуациях варпы в течение значительного числа тактов повторно исполняют одни и те же инструкции проверки условия, не продвигая выполнение программы. Это приводит к нерациональному использованию вычислительных ресурсов, увеличению нагрузки на аппаратный планировщик и снижению общей эффективности исполнения, поскольку ресурсы продолжают выделяться варпам, находящимся в состоянии циклического ожидания.

Для снижения отрицательного влияния активного ожидания применяется алгоритм Back-Off Warp Spinning (BOWS), предназначенный для выявления варпов, не продвигающихся по выполнению, и их временного исключения из процесса планирования [12]. Обнаружение таких варпов может осуществляться статически — на основе аннотаций или служебной информации, формируемой компилятором, — либо динамически, посредством аппаратурного контроля повторяющегося исполнения одной и той же инструкции на протяжении заданного числа тактов без изменения состояния выполнения. После идентификации соответствующего варпа его выполнение временно откладывается и исключается из основной очереди планирования на фиксированный или адаптивно определяемый интервал. По завершении данного интервала варп снова включается в число кандидатов на выполнение. Использование BOWS позволяет уменьшить долю вычислительных ресурсов, расходуемых на обслуживание циклов активного ожидания, и повысить эффективность планирования за счёт перераспределения исполнительных слотов в пользу варпов, выполняющих полезную работу. Вместе с тем практическое использование данного подхода требует дополнительной аппаратной поддержки, включая средства контроля продвижения выполнения и логику, управляющую повторным допуском варпов к исполнению после паузы.

6. Сводное сравнение и выводы

Таблица 1 обобщает рассмотренные ранее политики по аппаратной стоимости, ожидаемому эффекту и области применимости.

Таблица 1. Сравнение алгоритмов планирования варпов

Алгоритм	Аппаратные требования	Преимущества	Недостатки	Особенности применения
SRR	Минимальные	Простота, предсказуемость	Неэффективен при неоднородных нагрузках	Базовый эталонный метод
LRR	Только проверка готовности	Эффективнее, чем SRR	Возможна неравномерность задержек	Применяется в коммерческих ГП
GTO	Счётчик времени ожидания	Высокая локальность	Долгое ожидание исполнения для отдельных варпов	Базовый эталонный метод
TLS	Группировщик варпов	Меньше простоев, высокая локальность	Высокая аппаратная сложность	Эффективен на неоднородных задачах
SWL	Маска или ограничение активных варпов	Снижает конкуренцию за кэш	Сложность нахождения порога для всей программы	Используется при избытке активных потоков
CCWS	Модуль анализа промахов кэша	Адаптивность, меньше конфликтов за кэш	Высокая аппаратная сложность	При высокой кэш-нагрузке
DAWS	Предсказатель дивергенции, анализ кода	Предотвращает переполнение кэша	Сложный анализ, возможная неточность	Для программ с нерегулярным доступом к памяти
OAWS	Счётчик MSHR и модуль предсказания загрузки	Предотвращает простои из-за занятости ресурсов	Прогноз может быть неточен	При множественных обращениях к памяти
BOWS	Счётчики повторений и задержек	Исключает циклы активного ожидания	Сложная реализация блока обнаружения	Для нагрузок с синхронизацией потоков

Заключение

К настоящему времени не существует универсального алгоритма планирования варпов, применимого ко всем типам задач. В современных графических процессорах используются сочетания рассмотренных в работе подходов, которые выбираются с учетом нагрузки и аппаратных возможностей ГП. Исследования более сложных методов показывают, что значимого прироста производительности ГП можно достичь, если алгоритм учитывает характер доступа к данным. Однако издержки их реализации – серьезный ограничивающий фактор, поэтому внедрение таких идей происходит постепенно.

Ожидается появление гибридных алгоритмов, которые смогут динамически переключать алгоритм планирования в зависимости от наблюдаемого поведения нагрузки. Перспективным направлением также является программно-аппаратное взаимодействие: компилятор может передавать планировщику информацию о потенциальных проблемах, связанных с кэш-подсистемой или синхронизацией, а аппаратный блок – оперативно использовать эту информацию при выборе политики планирования.

Литература

1. E. Lindholm, J. Nickolls, S. Oberman, and J. Montrym, “NVIDIA Tesla: A Unified Graphics and Computing Architecture,” IEEE Micro, vol. 28, no. 2, pp. 39–55, Mar. 2008, doi: 10.1109/MM.2008.31.
2. S. Sandokji, F. Essa, and M. Fadel, “A survey of techniques for warp scheduling in GPUs,” in 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems

(ICICIS), Cairo, Abbassia, Egypt: IEEE, Dec. 2015, pp. 600–606. doi: 10.1109/IntelCIS.2015.7397284.
3. V. Narasiman, M. Shebanow, C. J. Lee, R. Miftakhutdinov, O. Mutlu, and Y. N. Patt, “Improving GPU performance via large warps and two-level warp scheduling,” in 2011 44th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), Dec. 2011, pp. 308–317. Accessed: Dec. 07, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/7851481
4. Jeong, S. Jang, and S. Hong, “Tournament Warp Scheduling: A Dynamic Policy Selector Exploiting Sub-cores in GPUs,” in 2025 International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC), Jul. 2025, pp. 1–5. doi: 10.1109/ITC-CSCC66376.2025.11137781.
5. “Fermi: NVIDIA’s Next Generation CUDA Compute Architecture.” [Online]. Available: https://www.nvidia.com/content/pdf/fermi_white_papers/nvidia_fermi_compute_architecture_whitepaper.pdf
6. J. Fang, Z. Wei, and H. Yang, “Locality-Based Cache Management and Warp Scheduling for Reducing Cache Contention in GPU,” Micromachines, vol. 12, no. 10, p. 1262, Oct. 2021, doi: 10.3390/mi12101262.
7. X. Mei and X. Chu, “Dissecting GPU Memory Hierarchy Through Microbenchmarking,” IEEE Trans. Parallel Distrib. Syst., vol. 28, no. 1, pp. 72–86, Jan. 2017, doi: 10.1109/TPDS.2016.2549523.
8. V. T. Vo and C. H. Kim, “Architecture exploration of recent GPUs to analyze the efficiency of hardware resources,” Bull. Electr. Eng. Inform., vol. 10, no. 2, pp. 917–926, Apr. 2021, doi: 10.11591/eei.v10i2.2736.
9. S. Liu, J. Ma, C. Zhao, X. Wan, and W. Wu, “LFWS: Long-Operation First Warp Scheduling Algorithm to Effectively Hide the Latency for GPUs,” IEICE Trans. Fundam. Electron. Commun. Comput. Sci., vol. E106.A, no. 8, pp. 1043–1050, Aug. 2023, doi: 10.1587/transfun.2022EAP1084.
10. T. G. Rogers, M. O’Connor, and T. M. Aamodt, “Cache-Conscious Wavefront Scheduling,” in 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture, Dec. 2012, pp. 72–83. doi: 10.1109/MICRO.2012.16.
11. B. Wang, Y. Zhu, and W. Yu, “OAWS: Memory Occlusion Aware Warp Scheduling,” in 2016 International Conference on Parallel Architecture and Compilation Techniques (PACT), Sep. 2016, pp. 45–55. doi: 10.1145/2967938.2967947.
12. A. ElTantawy and T. M. Aamodt, “Warp Scheduling for Fine-Grained Synchronization,” in 2018 IEEE International Symposium on High Performance Computer Architecture (HPCA), Feb. 2018, pp. 375–388. doi: 10.1109/HPCA.2018.00040.
13. T. G. Rogers, M. O’Connor, and T. M. Aamodt, “Divergence-Aware Warp Scheduling,” in 2013 46th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), Dec. 2013, pp. 99–110. Accessed: Dec. 07, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/7847618

WARP SCHEDULING ALGORITHMS IN GRAPHICS PROCESSORS

Titaev D.V.[1], Bocharov Y.I.[1], Zvyagin A.A.[1], Kovsharov I.D.[1], Butuzov V.A.[1]

[1] National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)

This paper presents the results of an analysis of contemporary hardware scheduling methods for groups of synchronous threads (warps) and their applicability to graphics processing units (GPUs). A classification and comparative assessment of current warp scheduling algorithms are proposed. For each approach, the operating principle, key advantages and limitations, as well as typical application scenarios, are briefly outlined. It is shown that the efficiency of these algorithms is determined both by workload characteristics and by the features of the memory subsystem.

МЕТОДЫ ОЦЕНИВАНИЯ РАБОТ ОБУЧАЮЩИХСЯ С ПРИМЕНЕНИЕМ СЕТЕЙ-ТРАНСФОРМЕРОВ

доц. к.т.н. Буряченко В.В.

Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева

В работе проведен обзор методов оценки ответов обучающихся на открытые вопросы. Для задач, которые требуют анализа естественного языка применяются различные подходы от построения векторных моделей текста, разибения слов на униграммы, биграммы и применения рекуррентных сетей. Современные подходы включат в себя использование сетей-трансформеров, которые основаны на применении механизма внимания, а также глубокое понимание контекста, что позволило сильно повысить качество анализа последовательностей и решать задачи генерации текста. Был проведен обзор существующих моделей и выполнен сбор собственного набора данных ответов обучающихся на русском языке. Для повышения качество обработки текста предложено дополнительно подавать на вход моделей информацию о видах дисциплины и типах вопроса, а также о средних оценках обучающихся. Экспериментальные исследования показали высокое качество современных моделей кодировщика BERT и GPT с тонкой настройкой при решении задач оценивания работ обучающихся. Точность по метрике F1 повысилась в среднем до 94% что превышает на 3-10% результаты моделей без дообучения.

Введение

Применение систем искусственного интеллекта в обучении студентов является широко обсуждаемой темой в современном обществе. В связи с развитием методов анализа текстовой информации появляются такие возможности как генерация тестовых заданий обучающимся, автоматизированное создание комплексов лабораторных работ, особенно в сферах, связанных с программированием и написанием программного кода. Системы искусственного интеллекта, основанные на больших лингвистических моделях и нейронных сетях с механизмом внимания способны в том числе являться ассистентами в усвоении материала и подсказывать недостающую информацию обучающимся.

Одной из наиболее интересных задач, которая может существенно облегчить процесс обучения, является автоматизация проверки работ обучающихся. В отличие от тестовых задач проверить ответов на открытые вопросы можно либо вручную преподавателем, либо с применением интеллектуальных систем. При этом целесообразным является подстройка системы на основе экспертных знаний преподавателей в данной области, что возможно за счёт дообучения используемой модели.

Обзор литературы

В качестве основного подхода в задачах анализа текстовой информации (Natural Language Processing, NLP) сейчас применяются нейронные сети, обученные предварительно на больших массивах данных. Затем, в зависимости от стоящей задачи: классификация настроений, отзывов пользователей, либо прогнозирование и построение вопросно-ответных систем, данная сеть дообучается на меньшем наборе данных, подобранном для конкретной задачи. В этой сфере применяются различные подходы, начиная с предобработки текстовых данных с целью нормализации и векторизации текстов, до глубоких нейронных сетей различных архитектур, основными из которых являются ячейки рекуррентных сетей LSTM и GRU, которые были наиболее перспективным подходом до 2017 года, так и более новые модели трансформеров и большие лингвистические модели (LLM), очень популярные сейчас.

В работе [1] авторы сравнивают возможности моделей трансформеров и фундаментальные модели в задаче оценки отзывов студентов, которые применяются для повышения качества образования. Авторы сравнили три модели трансформера и три фундаментальные модели на большом наборе данных эквадорского университета. Согласно экспериментам, модели трансформеры показывают лучшие результаты, а сеть DistilBERT заняла первое место с точностью оценки 84,90%.

В [2] авторы оценивают производительность сетей с механизмом внимания в задачах, где требуется учитывать конкретные сущности, например определения или понятия процедурных языков. Несмотря на то, что кодировщики, основанные на самовнимании, такие как BERT и GPT отлично справляются с задачами понимания естественных языков, в вопросах, связанных с конкретными терминами и понятиями, они часто показывают низкие результаты. Авторы предлагают повышать внимание к конкретным входным данным добавляя специальные токены

на отдельный вход сети для повышения внимания к определенным понятиям, что значительно повышает качество обработки данных с учетом контекста.

Сети-трансформеры и большие языковые модели хорошо справляются с анализом текстовых данных в области выявления настроения и классификации принадлежности текста, но понимание процессов принятия решений данными моделями является сложной задачей. Авторы [3] рассматривают работу различных моделей на наборах данных, таких как MAMS, SemEval, Naver с целью определения того, как эти модели делают свои предсказания. С применением тонкой настройки и дообучения модель RoBERTa достигла наилучших показателей в 87,2% на наборах MAMS и SemEval.

В работе [4] авторы выполняли сравнение нескольких моделей на наборе данных DT-Grade и показали, что точность классификации повысилась на 8-15% по сравнению с использованием рекуррентных нейронных сетей LSTM и достигла 88% по метрике F1. Авторы выполняли дообучение моделей T5, XLNET, BERT, DistilBERT, ALBERT и RoBERT, в задаче оценки ответов на открытые вопросы.

Модели трансформеры показали наилучшие результаты для большинства задач обработки текстовых данных по сравнению с классическими методами и рекуррентными сетями. Основными их отличиями от других методов можно назвать применение модифицированного механизма внимания с привлечением контекста и метаданных, а также использование контекстно-зависимых эмбеддингов [5, 6]. Архитектуры трансформеров включаются два модуля: кодировщик и декодировщик, каждый из которых можно использовать отдельно, за счёт чего их можно применять при решении различных классов задач, таких как генерация текста и классификация входных данных.

При подаче текста на выход сети выполняется его нормализация и токенизация. Токены формируются в виде словаря, и для поступающего на обученную модель выбирается номер слова в словаре либо количество раз, которое это слова встречается в тексте. Таким образом сеть может определить вид предложения путем сравнения векторизованных предложений. Словари современных моделей могут включать в себя несколько десятков тысяч слов, большие лингвистические модели также выполняют разбиение слова на подслова. В базовой модели BERT словарь включает тридцать тысяч слов [6]. Таким образом, на вход модели подаётся последовательность индексов токенов. BERT представляет собой модель кодировщика трансформера, поэтому лучше справляется с задачами понимания смысла слов, в том числе и многозначных, и задачами классификации входных данных.

В работе будет рассмотрено применение рассмотренных методов с использованием тонкой настройки сети для обучения модели в задаче оценки текстовых работ обучающихся по техническим дисциплинам. Возможность дообучения модели на собственном размеченном наборе данных может повысить точность предсказания оценки системой. Также авторы предлагают выполнять предварительную классификацию заданий по типу, что позволит повысить внимание на отдельные понятия и уточнить тип ответа, который преподаватель стремится получить от студента (например различия в вопросах-рассуждениях или определениях) для повышения точности оценки модели. Помимо этого, применение дообученной модели позволяет не только оценить качество ответа обучающегося, но и сформировать рекомендации для улучшения понимания предметной области.

Структура сети-трансформера для оценки работ обучающихся

Предлагаемый метод оценки работ обучающихся включает два вида вопросов такие, как:

– определения – ответ должен быть полным и соответствовать эталонному определению, но не должен включать в себя излишнюю информацию;

– развернутый ответ – данный ответ может содержать в себе как определение, так и дополнительную информацию, например, примеры использования той или иной технологии, свойства, характеристики понятий.

Информация о виде вопросов подается на один из входов сети для повышения внимания к используемому речевому аппарату и длине ответа (рис. 1).

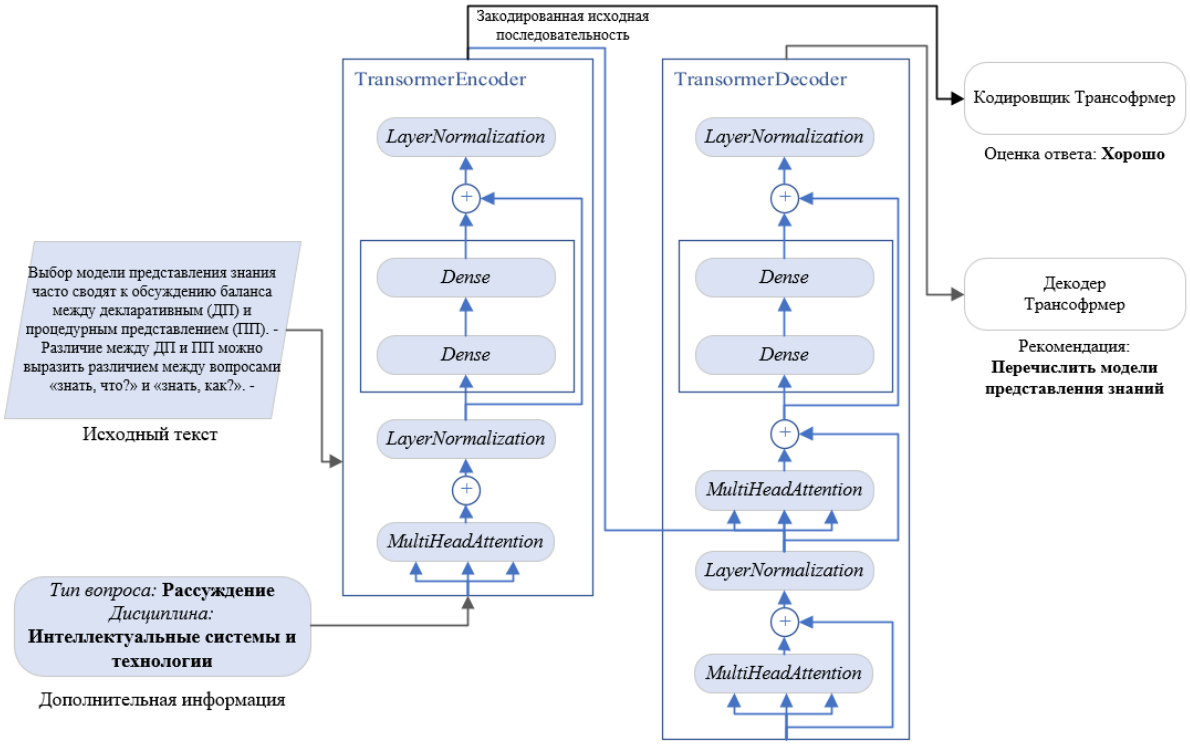


Рис. 1. Структура модели и данные, подаваемые на вход и получаемые на выходе

Для обработки данным используются классические подходы векторизации и нормализации текста, после чего полученный вектор подается на вход кодировщика. Текст ответа обучающегося дополняется информацией о типе вопроса, название дисциплины и дополнительными данными, например средним баллом обучающегося. После чего эта информация векторизуется и выполняется классификация отзыва на один из четырех классов оценки от «отличного» до «неудовлетворительного».

Для задачи классификации используется кросс-энтропийная функция потерь:

$$L = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log\left(\hat{y}_{i,c}\right), \quad (1)$$

где N – количество обучающих примеров, C – количество классов, $y_{i,c}$ – истинная оценка принадлежности i-го примера классу с, $\hat{y}_{i,c}$ – предсказанная моделью оценка принадлежности i-го примера классу c.

Декодировщик служит для формирования ответа системы на запрос пользователя: в данном случае он формирует рекомендацию обучающемуся для улучшения его ответа. Для генерации ответа используется функция ошибки на основе кросс-энтропии, которая учитывает тонкие параметры данной задачи. Для повышения внимания к типу вопроса и учету образовательного контекста используется функция активации на выходном слое, в которую подаются дополнительные параметры m_i, m_j:

$$a_{i,j} = \text{softmax}\left(\frac{Q_i K_j^T + f(m_i, m_j)}{\sqrt{d_k}}\right), \quad (2)$$

где m_i – метаданные о типе вопроса, m_j – название дисциплины, K_j^T – средняя оценка обучающегося Т по j-й дисциплине, d_k – размерности векторов Q и K для каждого слоя внимания.

Таким образом, одной из ключевых особенностей предлагаемого подхода является тонкая настройка обучающей модели трансформера с целью учета как особенностей обучающегося при наличии информации о его предыдущих ответах, так и конкретной дисциплины. Это дает возможность сети сгенерировать более точную оценку ответа и предложить рекомендации по улучшению работы студента при проведении повторной оценки уровня знаний. Для того, чтобы

система учитывала указанные параметры, в обучающем наборе данных должна быть выполнена разметка на классы ответов, а также указана принадлежность ответа к учебной дисциплине.

Экспериментальные исследования

Для проведения исследований использовались как известные набор данных SCoRE [7] и ASAP Kaggle Competition [8] для анализа, а также собственный набор данных на русском языке, включающий более 500 ответов обучающихся по четырем дисциплинам: «Информационные технологии», «Цифровая экономика», «Интеллектуальные системы и технологии», «Разработка и анализ требований». Длина ответов может быть различной от одного предложения до 10 (10 – 100 слов), что также добавляет сложностей в процесс обучения. Для тонкой настройки набор данных был разделен на обучающие, валидационные и тестовые данные в размере 60%, 20% и 20% соответственно. Каждый отзыв был вручную оценен экспертами на один из четырех классов: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Дополнительной сложностью является то, что в наборе изначально было менее 10% удовлетворительных и неудовлетворительных ответов, поэтому некоторые из этих данных были сгенерированы перемешиванием ответов и написанием случайных текстов. Описание сформированного набора данных представлено в таблице 1.

Таблица 1. Пример оценки качества ответа обучающегося

Вид предмета	**Количество ответов**	**Длина ответа, символов: Средняя, Максимальная**	**Количество сгенерированных ответов**	**Сложность используемого лексикона**	**Размер словаря**
Информационные технологии в цифровой экономике	621	358/1211	250	Низкая	2200
Разработка и анализ требований	254	259/655	150	Средняя	1800
Интеллектуальные системы и технологии	315	415/876	200	Высокая	1600
Теория автоматов и формальных языков	188	358/911	100	Высокая	1200

Для каждого предложения текст был объединен с соответствующим аспектом с помощью соответствующего маркера, обозначенного предварительно обученными моделями (например, [SEP] для BERT). Эта лексема эффективно отделяет основной текст от аспекта, позволяя модели четко понимать связь между двумя частями во время обучения. Использование этих лексем-разделителей позволяет моделям-трансформаторам обрабатывать входные данные как пару последовательностей, а не как одну конкатенированную строку, что позволяет предварительно обученным моделям концентрироваться на аспектах отдельных последовательностей, повышая точность прогнозирования. Обучающие данные в виде вопроса и ответа для BERT разбиваются на 512 токенов. Для задачи классификации ответов был добавлен выходной слой Softmax декодировщика, и модель дообучена на соответствие каждого из исходных предложений одному из четырех классов.

Пример оценки одного из ответов обучающихся на вопрос «Дайте определение виртуальных валют в контексте блокчейна» представлен в таблице 2.

В таблице 3 представлены результаты оценки качества предварительно обученных моделей трансформеров на наборах данных SCoRE и моделей, для которых была выполнена тонкая настройка на собственном наборе данных. Результаты значительно отличаются, и тонкая настройка сильно повышает качество предсказаний модели, что связано в первую очередь с дообучением сети на русскоязычном наборе данных.

Таблица 2. Пример оценки качества ответа обучающегося

Ответ обучающегося	**Вид вопроса и дисциплина**	**Рекомендация**	**Экспертная оценка**	**Итоговая оценка**
Виртуальная валюта – это денежные средства, не имеющие материального воплощения, которые могут использоваться как полноценный денежный знак. Блокчейн предлагает инновационные возможности для аутентификации и авторизации в мире цифровых технологий, исключающие необходимость использования централизованных систем. Благодаря своим особенностям технология Блокчейн находит все больше применений: авторское право, краудфайндинг инициативы, социальный авторитет, страхо-вание, реклама	Определение, ИТвЦЭ	Сократить ответ, указав только ключевые аспекты понятия виртуальных валют	Хорошо	Удовл.

Из полученных результатов можно сделать вывод, что модель BERT и модели класса GPT отлично справляются с классификацией ответов обучающихся на английском языке, а для использования данных моделей в образовательном процессе желательно выполнять дообучение на русскоязычных наборах данных. Рекуррентные модели, а также глубокие модели такие как DeepPavlov лучше работают с длинными последовательностями, что может не относиться к ответам студентов и эталонным ответам на вопросы.

Таблица 3. Сравнительная оценка качества оценки ответов обучающихся на наборах данных SCoRE и собственном наборе

Модель	**Точность**		**Полнота**		**F1-мера**	
Набор данных	**ScORE**	**Ours**	**ScORE**	**Ours**	**ScORE**	**Ours**
bert-base	0,85	0,81	0,89	0,82	0,87	0,81
bert-casedForDisc	**0,84**	**0,89**	**0,91**	**0,93**	**0,89**	**0,94**
Bi-LSTM-base	0,76	0,76	0,71	0,68	0,72	0,71
Bi-LSTM-casedForDisc	0,77	0,79	0,73	0,71	0,74	0,72
DeepPavlov-base	0,71	0,66	0,69	0,65	0,68	0,66
DeepPavlov-casedForDisc	0,72	0,68	0,67	0,63	0,69	0,71
Roberta-base	**0,87**	**0,83**	**0,88**	**0,86**	**0,88**	**0,87**
Roberta-casedForDisc	0,89	0,92	0,87	0,91	0,89	0,92
GPT-4-base	0,74	0,79	0,73	0,81	0,72	0,83

Обучение проводилось с использованием платформы Huggingface.co, которая предоставляет набор предобученных моделей, и собственных ресурсов настольного компьютера с видеокартой Nvidia RTX 3070 Ti. Дообучение модели на наборах данных SCoRE и Kaggle ASAP занимает не более 5 часов, а настройка на собственном наборе данных в связи с его небольшим размером потребовало менее 120 минут для одной модели.

Заключение

Сети-трансформеры, являющиеся основой для построения больших языковых моделей, позволяют решить многие задачи по обработке текстовых данных. Они заменили популярные ранее подходы, такие как использование рекуррентных сетей. В работе проводилось сравнение моделей в задаче анализа ответов обучающихся на вопросы с открытым ответом с целью их автоматической оценки. Для дообучения моделей были использованы классические наборы данных ScORE и ASAP, а также собран собственный набор ответов на русском языке по техническим дисциплинам. Благодаря тонкой настройке моделей с использованием собственного датасета качество оценки ответов повысилось до 94% по метрике F1, при этом наилучшие результаты показали модели BERT и GPT. Эксперименты показали высокую точность оценки работ по техническим дисциплинам, а также возможность формирования

отзывов в виде подсказок по улучшению ответа студента. Таким образом, большие языковые модели с применением дообучения можно использовать в сфере образования для проверки работ обучающихся при наличии контроля со стороны преподавателя и допустимости применения искусственного интеллекта в образовательном учреждении. Ключевым результатом можно представить предложенную модель трансформера на базе кодировщика BERT для классификации ответов обучающихся, а также разработанный набор текстовых данных по нескольким дисциплинам, который был использован для дообучения моделей.

Литература

1. I. Vega, J. Valencia, A. Arcos, D. Navarrete and M. Baldeon-Calisto, A Comparison Between Transformers and Foundation Models in Sentiment Analysis of Student Evaluation of Teaching, 2024 12th International Symposium on Digital Forensics and Security (ISDFS), San Antonio, TX, USA, 2024, pp. 1-7, doi: 10.1109/ISDFS60797.2024.10527264.
2. A. Gupta, G. Durrett, Effective Use of Transformer Networks for Entity Tracking, Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, pages 759–769, 2019.
3. I. Perikos, A. Diamantopoulos, Explainable Aspect-Based Sentiment Analysis Using Transformer Models. Big Data Cogn. Comput. 2024, 8, 141. https://doi.org/10.3390/bdcc8110141
4. A. Khayi Nisrine, R. Vasile, T. Lasang. Towards Improving Open Student Answer Assessment using Pretrained Transformers. The International FLAIRS Conference Proceedings. 34, pp. 1-4, 2021, doi: 10.32473/flairs.v34i1.128483.
5. Vaswani, A. et al. (2017). *Attention Is All You Need*. Advances in Neural Information Processing Systems (NIPS).
6. Devlin, J. et al. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. NAACL-HLT.
7. Dzikovska, M. , Nielsen, R. , Brew, C. , Leacock, C. , Giampiccolo, D. , Bentivogli, L. , Clark, P. , Dagan, I. and Dang, H. (2013), SemEval-2013 Task 7: The Joint Student Response Analysis and 8th Recognizing Textual Entailment Challenge, Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013).
8. The Hewlett Foundation: Short Answer Scoring, URL: https://www.kaggle.com/c/asap-sas, доступ получен 12.01.2025.

METHODS OF ASSESSING STUDENTS WORKS WITH TRANSFORMER NETWORKS

Buryachenko V.V.

Reshetnev Siberian State University of Science and Technology

The paper reviews methods for evaluating students answers to open-ended questions. For tasks that require natural language analysis, various approaches ranging from building vector models of text, unigrams, bigrams, and recurrent networks are used. Modern approaches include the use of transformer networks, which are based on the application of the attention mechanism, as well as a deep understanding of the context, which has greatly improved the quality of sequence analysis and text generation tasks. The existing models were reviewed and our own dataset of students' responses in Russian language was collected. In order to improve the quality of text processing, it was proposed to additionally feed the input of the models with information about the types of discipline and question types, as well as the average grades of students. Experimental studies have shown high quality of modern BERT encoder and GPT models with fine-tuning when solving problems of students' work evaluation. The accuracy on the F1 metric increased to 94% on average, which exceeds the results of models without additional training by 3-10%.

СРАВНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ ВЫСОКОТОЧНЫХ НАБОРОВ ДАННЫХ ДЛЯ ОЦЕНКИ АЛГОРИТМОВ ЛОКАЛИЗАЦИИ

студ. Могулев М., преп. Кирнос В.П., проф. Приоров А.Л.

Ярославский государственный университет им. П.Г. Демидова

В статье представлены результаты сравнительного анализа тестовых наборов RTS-GT с высокоточными эталонными траекториями для оценки алгоритмов одновременной локализации и построения карты пространства. Исследование проведено с использованием программной среды ROS2 на базе наборов данных, собранных 24 февраля и 7 марта 2022 года. Первый тестовый набор характеризуется протяженностью траектории 1,58 километра в условиях ясной погоды, второй имеет длину 0,98 километра и собран в условиях ледяного дождя. Выполнено тестирование SLAM-алгоритмов с использованием эталонных траекторий роботизированных тахеометров с точностью до 4 миллиметров. Представлены результаты количественного сравнения точности локализации и построения карт для различных условий эксплуатации. Сформулированы рекомендации по выбору тестовых наборов для тестирования алгоритмов в зависимости от требуемых характеристик точности и робастности.

Введение

Одновременная локализация и построение карты пространства (Simultaneous Localization and Mapping, SLAM) представляет собой фундаментальную задачу мобильной робототехники, решение которой необходимо для обеспечения автономной навигации в неизвестных средах. Развитие технологий компьютерного зрения и лидарного сканирования привело к появлению множества алгоритмов SLAM, что обусловило необходимость объективных методов оценки их точности.

Традиционные тестовые наборы с эталонными траекториями, полученными посредством систем глобального позиционирования (GNSS), обладают ограниченной точностью порядка нескольких сантиметров, что недостаточно для оценки современных высокоточных алгоритмов. Альтернативным подходом является применение роботизированных тахеометров (Robotic Total Stations, RTS), обеспечивающих точность измерений до 4 миллиметров в реальных внешних условиях. Тестовый набор RTS-GT содержит эталонные траектории с синхронизированными измерениями от лидаров, камер и GNSS-приемников, что позволяет проводить комплексное тестирование различных типов SLAM-алгоритмов.

Качество эталонных траекторий существенно зависит от погодных условий, характеристик среды и протяженности маршрута. В связи с этим возникает необходимость экспериментальной проверки применимости тестовых наборов RTS-GT для тестирования конкретных типов SLAM-алгоритмов.

Цель настоящей работы — экспериментальное исследование двух тестовых наборов RTS-GT в программной среде ROS2 с количественной оценкой точности локализации SLAM-алгоритмов в зависимости от погодных условий и характеристик траектории. На основе полученных результатов формулируются практические рекомендации по применению тестовых наборов для тестирования алгоритмов одновременной локализации и построения карты пространства.

Описание тестовых наборов

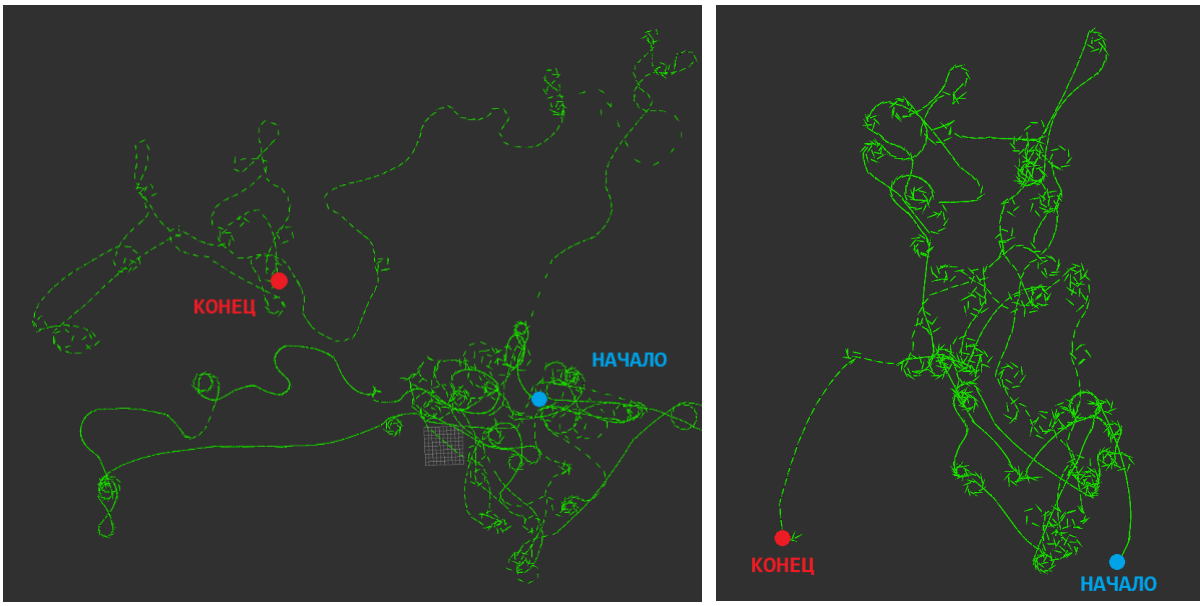
Исследуемые тестовые наборы получены с использованием роботизированных тахеометров Leica TS16 на мобильной платформе Clearpath Warthog. Сенсорное оборудование включало 64-лучевой лидар Ouster OS1-64, стереокамеры Intel RealSense D455, инерциальное измерительное устройство и двухчастотные GNSS-приемники. Система отслеживания с отражающими призмами обеспечивала измерение шестистепенной позы платформы с точностью до 4 миллиметров.

Таблица 1. Сравнительные характеристики исследуемых тестовых наборов RTS-GT

Параметр	24.02.2022	07.03.2022
Протяженность траектории, км	1,58	0,98
Длительность, мин:сек	18:42	11:28

Погодные условия	Ясная погода	Ледяной дождь
Средняя дальность RTS, м	39	20
Медианная точность RTS, мм	6,5	2,4
Медианная точность GNSS, мм	5,5	5,2
Ошибка трансляции, мм	15,2	28,7
Тип среды	Открытое пространство	Университетский кампус
Количество измерений RTS	11 тыс.	6,9 тыс.

Оба тестовых набора предоставляют данные в формате rosbag с полной синхронизацией всех сенсорных каналов, что обеспечивает возможность их непосредственного использования для тестирования SLAM-алгоритмов без дополнительной обработки. Пройденные пути из обоих тестовых наборов были визуализированы и представлены на рисунках 1а и 1б.



а) б)

Рисунок 1 - Визуализация пройденного пути в а) набор от 24.02.2022 и б) набор от 07.03.2022

Экспериментальная методология

Экспериментальное исследование проведено с использованием Robot Operating System 2 (ROS2) — современного фреймворка для разработки робототехнических приложений. Обработка тестовых наборов выполнялась на вычислительной системе под управлением Ubuntu 22.04 LTS. Визуализация траекторий осуществлялась посредством пакетов RViz2 и Plotjuggler. Оригинальные тестовые наборы конвертированы из формата ROS1 в ROS2 для обеспечения совместимости с современными SLAM-пакетами.

В рамках исследования проведено тестирование трех классов SLAM-алгоритмов:

1. Лидарные SLAM — методы на основе обработки облаков точек от 3D лидара с использованием алгоритмов сопоставления сканов (scan matching) и итеративного метода ближайших точек (ICP). Характеризуются высокой точностью в структурированных средах и устойчивостью к изменениям освещенности.

2. Визуальные SLAM — алгоритмы построения карт особых точек на основе данных стереокамер. Тестировались монокулярные и стереоскопические конфигурации для оценки влияния глубинной информации на точность локализации.

3. Мультисенсорные SLAM — методы слияния данных от лидара, камер и IMU с использованием фильтра расширенного Калмана (EKF) для повышения робастности локализации.

Для количественной оценки точности работы SLAM-алгоритмов использовались стандартные метрики сравнения траекторий относительно высокоточных эталонных измерений RTS.

Абсолютная ошибка траектории (Absolute Trajectory Error, ATE) определяет среднеквадратичное отклонение между оцененными и эталонными позициями робота:

$$ATE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{p}_i^{est} - \boldsymbol{p}_i^{gt}\right\|^2}$$

n — количество измерений. Метрика характеризует глобальную согласованность траектории и чувствительна к накоплению ошибок локализации.

Относительная ошибка позы (Relative Pose Error, RPE) оценивает локальную точность путем анализа изменения позы между последовательными временными интервалами:

$$RPE_{\Delta t} = \left\|\left(\boldsymbol{T}_i^{gt}\right)^{-1}\boldsymbol{T}_{i+\Delta t}^{gt} - \left(\boldsymbol{T}_i^{est}\right)^{-1}\boldsymbol{T}_{i+\Delta t}^{est}\right\|$$

Δt — временной интервал. RPE характеризует дрейф локализации на коротких участках траектории.

Для обоих тестовых наборов вычислялись медианные значения, первый и третий квартили распределения ошибок ATE и RPE. Дополнительно анализировалась стабильность измерений путем вычисления стандартного отклонения ошибок на скользящих окнах длиной 10 секунд. Статистическое сравнение результатов проводилось с использованием критерия Манна-Уитни для оценки значимости различий в точности локализации при изменении погодных условий.

Результаты экспериментального исследования

Анализ точности эталонных траекторий показал существенные различия между наборами. Набор от 7 марта продемонстрировал на 63% более высокую медианную точность RTS-измерений (2,4 мм против 6,5 мм), несмотря на неблагоприятные погодные условия. Это объясняется меньшей средней дальностью отслеживания (20 м против 39 м), что обеспечило более стабильное позиционирование отражающих призм. Точность GNSS-измерений оставалась сопоставимой для обоих тестовых наборов (5,5 мм и 5,2 мм соответственно), что подтверждает робастность спутниковой навигации к атмосферным осадкам в открытых пространствах. При этом ошибка трансляции для набора 07.03 оказалась на 89% выше (28,7 мм против 15,2 мм), что связано с более короткими базовыми линиями между призмами при меньшей протяженности траектории.

Тестирование на тестовом набор с ясной погодой и протяженностью 1,58 км выявило различную производительность методов. Результаты представлены в таблице 2.

Таблица 2. Результаты тестирования SLAM-алгоритмов на тестовом наборе от 24.02.2022

Алгоритм	**Тип**	**ATE, м**	**RPE, м**	**Станд. откл., м**
FAST-LIO2	Лидар	0,087	0,012	0,034
LIOSAM	Лидар	0,105	0,015	0,041
ORB-SLAM3	Стерео	0,142	0,023	0,067
VINS-Mono	Моно	0,198	0,031	0,089
R3LIVE	Лидар+IMU	0,073	0,009	0,028

Лидарные методы продемонстрировали ATE 0,087–0,105 м с высокой локальной согласованностью (RPE 0,009–0,015 м). Мультисенсорный подход показал наилучший результат (0,073 м) за счет эффективного слияния данных лидара и IMU. Визуальные SLAM показали значительно более высокие ошибки (0,142–0,198 м) из-за ограниченной текстуры поверхностей в зимних условиях. Стандартное отклонение для визуальных методов (0,067–0,089 м) превышает лидарные в 2–3 раза. Критические участки выявлены на поворотах, где визуальные методы демонстрировали пики ATE до 0,45 м.

Набор с ледяным дождем создал существенный вызов для всех типов алгоритмов. Результаты приведены в таблице 3.

Таблица 3. Результаты тестирования SLAM-алгоритмов на тестовом набор от 07.03.2022

Алгоритм	**Тип**	**ATE, м**	**RPE, м**	**Станд. откл., м**
FAST-LIO2	Лидар	0,094	0,014	0,038
LIOSAM	Лидар	0,118	0,017	0,049
ORB-SLAM3	Стерео	0,231	0,041	0,112
VINS-Mono	Моно	0,327	0,058	0,156
R3LIVE	Лидар+IMU	0,081	0,011	0,032

Визуальные методы показали критическую деградацию: ATE возросла на 63–65% (до 0,231–0,327 м), RPE увеличилась вдвое. Ледяной дождь на оптических поверхностях резко снизил количество детектируемых визуальных особенностей. Лидарные алгоритмы продемонстрировали робастность с увеличением ATE лишь на 8–12%, сохраняя стабильные измерения при осадках. Отражения от капель создавали шумовые точки, требующие агрессивной фильтрации. Мультисенсорный подход сохранил лучшую производительность (0,081 м) с увеличением ошибки всего на 11% за счет компенсации сбоев лидара данными IMU.

Сводное сравнение представлено в таблице 4.

Таблица 4. Сравнительный анализ производительности

Алгоритм	**ATE 24.02, м**	**ATE 07.03, м**	**Δ, %**	**Робастность**
FAST-LIO2	0,087	0,094	+8,0	Высокая
LIOSAM	0,105	0,118	+12,4	Высокая
ORB-SLAM3	0,142	0,231	+62,7	Низкая
VINS-Mono	0,198	0,327	+65,2	Низкая
R3LIVE	0,073	0,081	+11,0	Очень высокая

Критерий Манна-Уитни подтвердил значимость различий ($p < 0,01$) для визуальных методов между тестовыми наборами, в то время как для лидарных различия статистически незначимы ($p > 0,05$). Анализ распределения ошибок показал бимодальное распределение ATE для визуальных методов на наборе от 07.03 с пиками на участках интенсивных осадков. Лидарные алгоритмы сохранили нормальное распределение с незначительным смещением медианы.

Протяженность траектории влияет на накопление ошибок: для набора от 24.02 визуальные методы без замыкания циклов показали линейное возрастание RPE на интервалах 50–100 м. Более высокая точность эталона набора от 07.03 (2,4 мм) не привела к пропорциональному улучшению SLAM из-за доминирования сенсорных ошибок. Критические участки на поворотах >90° вызывали пики RPE до 0,08 м (лидарные) и 0,15 м (визуальные) из-за ограниченного перекрытия сканов.

Набор от 24.02.2022 оптимален для тестирования долгосрочной стабильности алгоритмов на протяженных траекториях. Длина маршрута 1,58 км обеспечивает условия для оценки накопления ошибок локализации и эффективности замыкания циклов. Стабильные погодные условия минимизируют внешние возмущения, что позволяет сосредоточиться на внутренних характеристиках алгоритмов. Рекомендуется для бенчмаркинга визуальных SLAM-методов и калибровки параметров.

Набор от 07.03.2022 предназначен для оценки робастности к неблагоприятным погодным факторам. Условия ледяного дождя создают реалистичный сценарий эксплуатации с ограниченной видимостью. Меньшая протяженность (0,98 км) компенсируется более высокой точностью эталона (2,4 мм), критичной для анализа локальных ошибок. Особенно ценен для тестирования лидарных и мультисенсорных методов для всепогодной эксплуатации.

При переводе из ROS1 в ROS2 необходимо обеспечить корректную трансформацию типов сообщений с сохранением временных меток микросекундной точности для синхронизации мультисенсорных данных.

Для лидарных методов рекомендуется фильтрация дальности 0,5–50 м. Визуальные SLAM требуют калибровки стереопары по эталонным параметрам из метаданных набора для минимизации систематических ошибок.

Для набора от 07.03 лидарные алгоритмы требуют статистической фильтрации (mean_k=50, stddev=1,0) для удаления шумовых точек от осадков. Визуальные методы выигрывают от адаптивной гистограммной эквализации (CLAHE) для компенсации неравномерности освещенности.

Тестирование проведено на наборах из университетского кампуса, что может ограничивать обобщаемость выводов на городские, лесные или промышленные среды. Не все современные методы на основе глубокого обучения включены в тестирование из-за отсутствия стабильных портов на ROS2. Временной промежуток между наборами (менее двух недель) ограничивает разнообразие сезонных условий.

Заключение

Проведено экспериментальное исследование набором RTS-GT для оценки SLAM-алгоритмов в среде ROS2. Тестирование лидарных, визуальных и мультисенсорных методов выявило ключевые факторы, влияющие на точность локализации в различных условиях.

Лидарные SLAM-алгоритмы демонстрируют высокую робастность к погоде с увеличением ATE не более 12% при переходе от ясных условий к осадкам. Визуальные методы показали деградацию производительности на 63–65% в неблагоприятных условиях. Мультисенсорные подходы обеспечили наилучшую комбинацию точности (ATE 0,073–0,081 м) и стабильности.

Набор от 24.02 рекомендуется для оценки долгосрочной стабильности и калибровки, набор от 07.03 — для тестирования робастности к погодным факторам. Высокая точность эталонных RTS-траекторий (2,4–6,5 мм) превосходит традиционные GNSS-набор более чем на порядок величины.

Практическая значимость заключается в предоставлении количественных данных о производительности SLAM-алгоритмов для обоснованного выбора методов локализации. Рекомендации по конфигурации в ROS2 могут быть непосредственно использованы при разработке систем автономной навигации.

Направления дальнейших исследований включают тестирование на наборах с различными сезонными условиями, оценку методов семантического SLAM и алгоритмов на основе глубокого обучения, а также анализ энергоэффективности для мобильных платформ с ограниченными ресурсами.

Литература

1. Vaidis M., Shahraji Z. T., Giguère P., Pomerleau F. RTS-GT: Robotic Total Stations Ground Truthing dataset // IEEE Robotics and Automation Letters. 2023. Vol. 8, No. 10. P. 6738–6745. DOI: 10.1109/LRA.2023.3313920.

2. Daum E., Pomerleau F. Benchmarking ground truth trajectories with robotic total stations // 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Detroit, MI, USA, 2023. P. 8710–8717. DOI: 10.1109/IROS55552.2023.10341737.

3. Xu W., Cai Y., He D., Lin J., Zhang F. FAST-LIO2: Fast Direct LiDAR-Inertial Odometry // IEEE Transactions on Robotics. 2022. Vol. 38, No. 4. P. 2053–2073. DOI: 10.1109/TRO.2022.3141876.

4. Shan T., Englot B., Meyers D., Wang W., Ratti C., Rus D. LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping // 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Las Vegas, NV, USA, 2020. P. 5135–5142. DOI: 10.1109/IROS45743.2020.9341176.

5. Campos C., Elvira R., Rodriguez J. J. G., Montiel J. M. M., Tardós J. D. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM // IEEE Transactions on Robotics. 2021. Vol. 37, No. 6. P. 1874–1890. DOI: 10.1109/TRO.2021.3075644.

6. Qin T., Li P., Shen S. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator // IEEE Transactions on Robotics. 2018. Vol. 34, No. 4. P. 1004–1020. DOI: 10.1109/TRO.2018.2853729.

7. Lin J., Zhang F. R3LIVE: A Robust, Real-time, RGB-colored, LiDAR-Inertial-Visual tightly-coupled state Estimation and mapping package // 2022 IEEE International Conference on Robotics and Automation (ICRA). Philadelphia, PA, USA, 2022. P. 10672–10678. DOI: 10.1109/ICRA46639.2022.9811935.

8. Chen X., Wang H., Li Y., Zhang G. Visual SLAM: What Are the Current Trends and What to Expect? // Sensors. 2022. Vol. 22, No. 23. Article 9297. DOI: 10.3390/s22239297.

9. Li Y., Brasch N., Wang Y., Navab N., Tombari F. A review of visual SLAM for robotics: evolution, properties, and future applications // Frontiers in Robotics and AI. 2024. Vol. 11. Article 1347985. DOI: 10.3389/frobt.2024.1347985.

10. Yarovoi A., Vasilyev I., Tsoy T., Shamsuarov A. Review of simultaneous localization and mapping (SLAM) techniques for construction applications // Automation in Construction. 2024. Vol. 160. Article 105306. DOI: 10.1016/j.autcon.2024.105306.

11. Pfreundschuh P., Kumar V., Zimmermann F., Reijgwart V., Dubé R., Siegwart R., Cramariuc A., Cadena C., Chli M. Benchmarking SLAM Algorithms in the Cloud: The SLAM Hive Benchmarking Suite // IEEE Robotics and Automation Letters. 2024. Vol. 9, No. 7. P. 6181–6188. DOI: 10.1109/LRA.2024.3402651.

12. Burnett K., Yoon D. J., Wu Y., Li A. Z., Zhang H., Lu S., Qian J., Tseng W.-K., Lambert A., Leung K. Y. K., Schoellig A. P., Barfoot T. D. YUTO MMS: A comprehensive SLAM dataset for urban mobile mapping with tilted LiDAR // Scientific Data. 2024. Vol. 11. Article 614. DOI: 10.1038/s41597-024-03452-3.

13. Liu Z., Wang Y., Chen X., Zhang H. Research on Positioning Accuracy of Mobile Robot in Indoor Environment // Sensors. 2023. Vol. 23, No. 23. Article 9603. DOI: 10.3390/s23239603.

14. Silva R., Santos V., Premebida C. Global localization for mobile robots in symmetrical indoor environments: a review, practical challenges, and experimental validation // Robotica. 2026. Vol. 44, No. 3. P. 1039–1054. DOI: 10.1017/S0263574724000079.

15. Geiger A., Lenz P., Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite // 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA, 2012. P. 3354–3361. DOI: 10.1109/CVPR.2012.6248074.

COMPARATIVE STUDY OF HIGH-PRECISION DATASETS FOR SLAM ALGORITHM EVALUATION

Mogulev M., Kirnos V.P., Priorov A.L.
P.G. Demidov Yaroslavl State University

The paper presents the results of a comparative analysis of RTS-GT datasets with high-precision ground truth trajectories for evaluating simultaneous localization and mapping algorithms. The study was conducted using the ROS2 software environment based on datasets collected on February 24 and March 7, 2022. The first dataset is characterized by a trajectory length of 1,58 kilometers under clear weather conditions, the second has a length of 0,98 kilometers and was collected under freezing rain conditions. SLAM algorithms were tested using ground truth trajectories from robotic total stations with an accuracy of up to 4 millimeters. The results of quantitative comparison of localization accuracy and map building for various operating conditions are presented. Recommendations are formulated for selecting datasets for algorithm testing depending on the required accuracy and robustness characteristics.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ СВЕРТОЧНЫХ И ГИБРИДНЫХ НЕЙРОСЕТЕВЫХ АРХИТЕКТУР ДЛЯ РАСПОЗНАВАНИЯ ВИДОВ МОДУЛЯЦИИ РАДИОСИГНАЛОВ В УСЛОВИЯХ МЕДЛЕННЫХ ЗАМИРАНИЙ

асп. Савин С.А., асп. Усков Н.А., доц. Красавин К.С.

Ярославский государственный университет им. П.Г. Демидова

Представлены результаты сравнительного анализа двух подходов к автоматическому распознаванию видов цифровой модуляции радиосигналов на основе глубокого обучения: сверточной нейронной сети (CNN_NET) и гибридной архитектуры, объединяющей сверточные слои и долгую краткосрочную память (CNN+LSTM). Обе модели обучались на датасете RadioML2018.01A с использованием аугментации данных, имитирующей медленные замирания Релея и Райса. Эксперименты проводились для восьми видов модуляции (OOK, 4ASK, BPSK, QPSK, 16QAM, AM-SSB-WC, FM, GMSK) при отношении сигнал/шум от –20 до 30 дБ с шагом 2 дБ. Показано, что сверточная нейронная сеть CNN_NET демонстрирует низкую эффективность: общая точность классификации составила лишь 57,04%, при этом для модуляции QPSK значение F1-score оказалось нулевым. В то же время гибридная архитектура CNN+LSTM, предложенная в работе [1], обеспечивает значительно более высокую точность (84,99%) за счет учета временных зависимостей в структуре сигнала. Полученные результаты подтверждают неэффективность простых сверточных архитектур для распознавания модуляций в условиях замираний и обосновывают необходимость рекуррентных компонентов.

Введение

Автоматическое распознавание вида модуляции (AMR) необходимо в системах когнитивного радио и программно-определяемого радио, а также в системах радиомониторинга. В реальных каналах связи сигналы подвержены медленным замираниям Релея (отсутствие прямой видимости) и Райса (наличие прямого луча), что вызывает плавные изменения амплитуды и фазы и усложняет идентификацию типа модуляции.

Сверточные нейронные сети (CHC, CNN) эффективно выделяют локальные признаки, но ограничены в моделировании долговременных зависимостей. Рекуррентные сети LSTM, напротив, учитывают контекст, но могут упускать локальные детали. Гибридные архитектуры CNN+LSTM объединяют преимущества обоих подходов. В работе [1] предложена гибридная модель, достигшая точности 84,99% на датасете RadioML2018.01A с аугментацией медленных замираний. Для обоснования преимущества гибридного подхода необходимо прямое сравнение с чисто сверточной архитектурой в идентичных условиях.

Цель работы – сравнительный анализ сверточной нейронной сети CNN_NET (собственная реализация) и гибридной сети CNN+LSTM [1] при распознавании восьми видов модуляции в условиях медленных замираний.

1 Описание датасета и предобработка данных

Для исследований использована открытая база данных RadioML2018.01A [2], содержащая I/Q-сигналы длиной 1024 отсчета. Выбраны восемь достаточно популярных видов модуляции: OOK, 4ASK, BPSK, QPSK, 16QAM, AM-SSB-WC, DM, GMSK. Данные модуляции охватывают широкий спектр сигнально-кодовых конструкций современных систем связи (терагерцовые системы [3], когерентные PON [4], сотовые системы 4G/5G, спутниковая связь [5], беспроводные сети Wi-Fi [6] и др.). Диапазон отношения сигнал-шум (SNR) от -20 до 30 дБ с шагом 2 дБ.

Для имитации медленных замираний к обучающей выборке применялась аугментация, аналогичная [1]. Коэффициент передачи канала с замираниями Релея моделировался как комплексная гауссовская случайная величина: $h_{\text{Релея}} = h_1 + \text{j}h_2$, где $h_1, h_2 \sim N(0; 0{,}5)$. Для замираний Райса использовалась модель с фактором $K = 2$: $h_{\text{Райса}} = \sqrt{\frac{K}{K+1}}e^{j\varphi} + \sqrt{\frac{1}{K+1}}(h_1 + \text{j}h_2)$, где $\varphi \sim U[0,2\pi]$. Коэффициент h постоянен в пределах одного пакета (условие медленных замираний). Аугментация (размножение данных) применялась к 80% обучающих примеров с равной вероятностью выбора типа замираний (смешанный тип). Мощность сигнала нормировалась для сохранения исходных SNR.

2 Архитектура нейронных сетей

2.1 Сверточная нейронная сеть CNN_NET

Архитектура СНС CNN_NET включает четыре последовательных блока. Входной тензор (2, 1024) обрабатывается:

– Блок 1: Conv1d (2→32, ядро 15, шаг 2)+BatchNorm+ReLU+Dropout(0,2);
– Блок 2: Conv1d (32→64, ядро 7, шаг 2)+BatchNorm+ReLU+Dropout(0,2);
– Блок 3: Conv1d(64→128, ядро 5, шаг 2)+BatchNorm+ReLU+Dropout(0,2);
– Блок 4: Conv1d(128→256, ядро 3, шаг 2)+BatchNorm+ReLU+Dropout(0,2);
– Global Adaptive Average Pooling (выход 256);
– Полносвязный слой: 256→128 (ReLU, Dropout 0,3)→8(линейный).

Число параметров ~ 650 тыс.

2.2 Гибридная сеть CNN+LSTM

Архитектура из [1] использует раздельную обработку I и Q компонент двумя ветвями сверточных слоев. Каждая ветвь содержит два сверточных слоя (64 фильтра 3x3 LeakyReLU), MaxPooling (3x3, шаг 2), затем два сверточных слоя (128 фильтров 3x3) MaxPooling. После объединения и операции Flatten данные подаются в LSTM-слой (256 скрытых единиц), затем на полносвязные слои с Dropout 0,5.

3 Условия экспериментов

Обучение обеих моделей проводилось с одинаковыми гиперпараметрами:

– размер батча = 32;
– количество эпох = 40;
– оптимизатор Adam (lr= 0,001);
– планировщик OneCycleLR (max_lr= 0,005);
– функция потерь – кросс-энтропия;
– аугментация замираний только на обучающей выборке.

Эксперименты выполнены на системе: Intel Xeon E5-2673 v3, 32 ГБ ОЗУ, NVIDIA GXT 1660 Super, Windows 10, PyTorch.

Метрики: Precision, Recall, F1-score, макро-усредненные (Macro avg) и взвешенные (Weighted avg) значения.

4 Полученные результаты и их анализ

4.1 Обучение моделей

Модель CNN_NET (рисунок 1) демонстрирует переобучение: тренировочные потери снижаются до уровня <0,1, тогда как валидационные колеблются с амплитудой до 27 (минимум 0,80 на 32-й эпохе). Тренировочная точность стабильна (0,7-0,8), но валидационная варьируется от 0,4 до 0,75, что указывает на низкую обобщающую способность и непригодность модели для практического применения.

Модель CNN+LSTM показывает сбалансированное обучение: тренировочные и валидационные потери сходятся к уровням ~0,1 и ~0,2 соответственно (разница 0,1–0,15). Точность обеих выборок стабильно достигает 0,85–0,9 с минимальным расхождением (~0,05), подтверждая высокую обобщающую способность и устойчивость к шуму.

Критерий успешного обучения – сходимость валидационных метрик к тренировочным без резких колебаний, что достигнуто только в гибридной архитектуре CNN+LSTM.

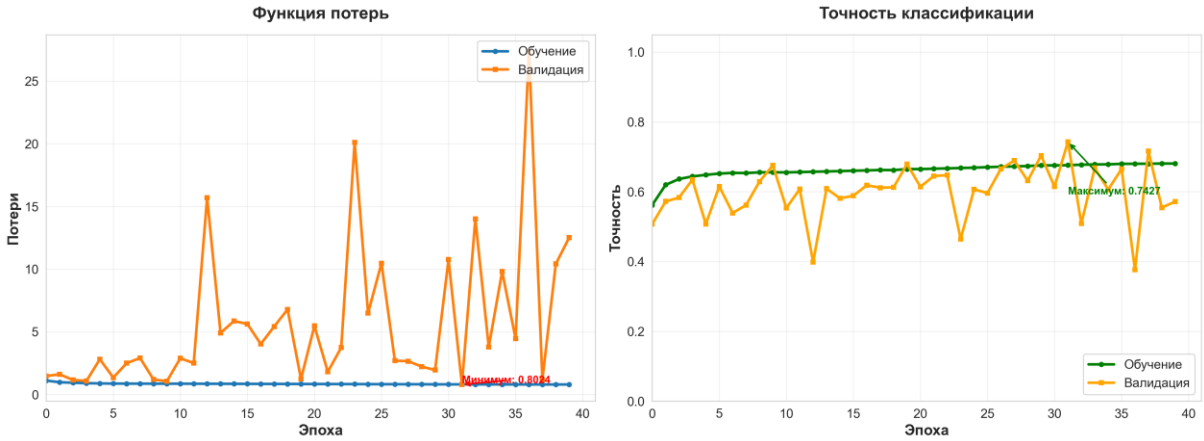


Рисунок 1 – Графики функции потерь и точности классификации сверточной нейронной сети CNN_NET

4.2 Сравнение метрик классификации

В таблицах 1 и 2 приведены результаты сравнения для СНС CNN_NET (по результатам тестирования) и гибридной нейронной сети CNN+LSTM (данные из [1]). Усредненные метрики сведены в таблице 3.

Таблица 1 – Результаты тестирования СНС CNN_NET

Модуляция	Precision	Recall	F1-score
OOK	0,502	0,686	0,580
4ASK	0,812	0,722	0,765
BPSK	1,000	0,599	0,749
QPSK	0,000	0,000	0,000
16QAM	0,263	0,923	0,409
AM-SSB-WC	1,000	0,652	0,789
FM	1,000	0,698	0,823
GMSK	1,000	0,281	0,439

Таблица 2 – Результаты тестирования гибридной НС CNN+LSTM (по данным [1])

Модуляция	Precision	Recall	F1-score
OOK	0,828	0,984	0,899
4ASK	0,947	1,000	0,973
BPSK	0,825	0,993	0,899
QPSK	0,817	0,804	0,810
16QAM	0,769	0,900	0,830
AM-SSB-WC	0,905	0,999	0,951
FM	0,838	0,899	0,866
GMSK	0,833	0,859	0,845

Таблица 3 – Усредненные метрики

Метрика	CNN_NET	CNN+LSTM[1]
Macro avg Precision	0,697	0,858
Macro avg Recall	0,571	0,905
Macro avg F1-score	0,569	0,872
Weighted avg F1-score	0,569	0,872
Общая точность (Accuracy)	57,04%	84,99%

Сеть CNN_NET показала крайне низкие результаты: для QPSK F1=0 (ни одного верно распознанного сигнала), для 16QAM и GMSK F1=0,409 и 0,439. Высокие Precision для BPSK, AM-SSB_WC, FM, GMSK обусловлены малым числом ложных срабатываний, но низкой полнотой. Гибридная сеть стабильно высока по всем классам (минимальный F1=0,810 для QPSK).

4.3 Анализ матриц ошибок

На рисунке 2 представлена матрица ошибок СНС CNN_NET. Сигналы QPSK ошибочно классифицированы как 16QAM (~53%), OOK (~28%) и др. Также заметно перепутывание FM и GMSK (~15%).

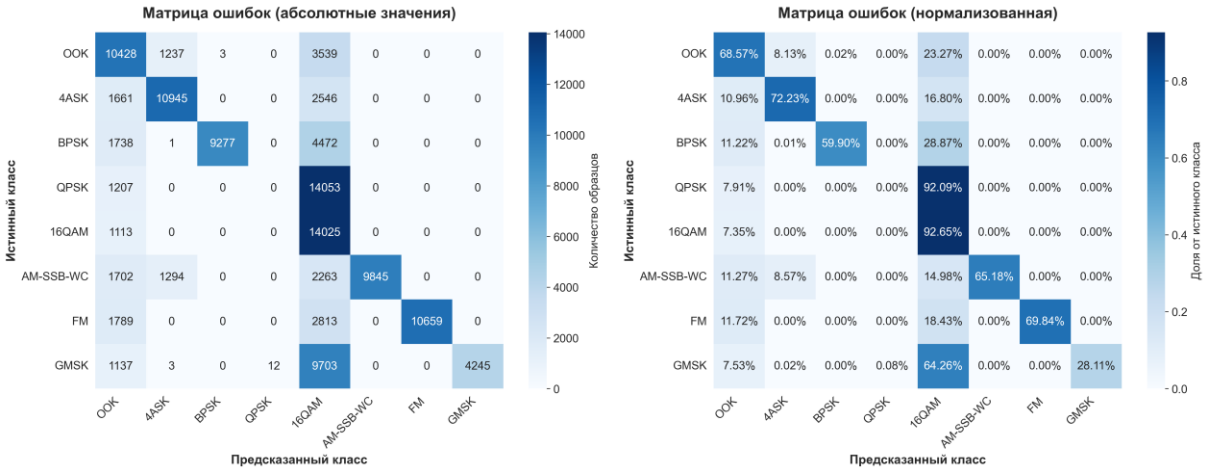


Рисунок 2 – Матрица ошибок сверточной нейронной сети CNN_NET

Матрица ошибок гибридной сети [1] показывает существенно меньшее перепутывание: ошибки между QPSK и 16QAM <10%, точность классификации FM и GMSK >85%.

Заключение

Сравнительный анализ показал, что сверточная сеть CNN_NET неэффективна для распознавания модуляций в условиях медленных замираний (общая точность 57,04%, для QPSK F1=0). Гибридная архитектура CNN+LSTM [1] превосходит ее по всем метрикам: точность выше на 27,95%, макро F1 – на 0,303. Наибольший выигрыш достигнут для модуляций, чувствительных к фазовым искажениям (QPSK, 16QAM, FM, GMSK). Результаты подтверждают необходимость учета временных зависимостей с помощью рекуррентных слоев для устойчивого распознавания в сложных каналах. Дальнейшие исследования могут включать оптимизацию гибридных архитектур и применение механизмов внимания.

Литература

1. Приоров А.Л., Ушакова А.М. Тимонькин М.В. Разработка нейросетевого алгоритма распознавания вида модуляции радиосигнала в условиях медленных замираний в радиоканале // Цифровая обработка сигналов. 2025. № 4. С. 11-15.
2. O'Shea T.J., Roy T., Clancy T.C. Over-the-Air Deep Learning Based Radio Signal Classification // IEEE J. Sel. Topics Signal Process. 2018. Vol. 12. No. 1. P. 168-179.
3. Jayaprasath Elumalai, et al. Free-space optical communications at 4 Gbit/s data rate with a terahertz laser // Communications physics. 2026. Vol. 9. Article 35. – ISSN 2399-3650. – DOI: 10.1038/s42005-025-02471-w. – URL: https://www.nature.com/articles/s42005-025-02471-w (дата обращения: 10.02.2026).
4. Shimada K., Inagaki M., Matsumoto R., Kodama T. IQ-channel multiplexed DP-4ASK signal using power loading for downlink coherent PON system with access-span length difference // Optics Communications. 2024. Vol. 569. Article 130837. – ISSN 0030-4018. – DOI: 10.1016/j.optcom.2024.130837. – URL: https://www.sciencedirect.com/science/article/abs/pii/S0030401824005741 (дата обращения: 10.02.2026).
5. Krikidis I. Quantum optical integrated sensing and communication with homodyne BPSK detection. – DOI 10.48550/arXiv.2601.07034. – Текст: электронный // arXiv.org: [сайт]. 2026. 6 p. – URL: https://arxiv.org/abs/2601.07034 (дата обращения: 11.02.2026). – Номер документа: arXiv:2601.07034v1.
6. Abdelrahman A.E., Younis M.B., Selim M.O. et al. A 1.54-pJ/b 64-Gb/s 16-QAM intradyne coherent optical receiver in 28-nm CMOS // IEEE Journal of Solid-State Circuits. 2025. Vol. 60, no. 12. P. 3334–3350. – ISSN 0018-9200. – DOI: 10.1109/JSSC.2025.3526795. – URL: https://ieeexplore.ieee.org/document/11277336 (дата обращения: 11.02.2026).
7. Беляева М.Б. Распознавание модуляции: что можно узнать, если заранее ничего не известно // Цифровая обработка сигналов. 2013. № 2. С. 55-63.
8. Аджемов С.С., Кленов Н.В., Терешонок М.В., Чиров Д.С. Методы распознавания видов цифровой модуляции сигналов в когнитивных радиосистемах // Вестник Московского университета. Серия 3. Физика. Астрономия. 2015. № 6. С. 19-27.
9. Крон Дж., Бейлевельд Г., Аглаэ Б. Глубокое обучение в картинках. Визуальный гид по искусственному интеллекту. СПб.: Питер. 2020. 400 с.

COMPARATIVE ANALYSIS OF THE EFFECTIVENESS OF CONVOLUTIONAL AND HYBRID NEURAL NETWORK ARCHITECTURES FOR RECOGNIZING TYPES OF RADIO SIGNAL MODULATION IN CONDITIONS OF SLOW FADING

Savin S.A., Uskov N.A., Krasavin K.S.

P.G. Demidov Yaroslavl State University

The results of a comparative analysis of two approaches to automatic recognition of digital radio signal modulation based on deep learning are presented: a convolutional neural network (CNN_NET) and a hybrid architecture combining convolutional layers and long-term short-term memory (CNN+LSTM). Both models were trained on the RadioML2018.01A dataset using data augmentation simulating Rayleigh and Rice slow fades. Experiments were performed for eight types of modulation (OOK, 4ASK, BPSK, QPSK, 16QAM, AM-SSB-WC, FM, GMSK) at a signal-to-noise ratio from -20 to 30 dB in 2 dB increments. It is shown that the CNN_NET convolutional neural network demonstrates low efficiency: the overall classification accuracy was only 57.04%, while the F1-score value for QPSK modulation turned out to be zero. At the same time, the CNN+LSTM hybrid architecture proposed in [1] provides significantly higher accuracy (84.99%) by taking into account time dependencies in the signal structure. The results obtained confirm the inefficiency of simple convolutional architectures for recognizing modulations under fading conditions and justify the need for recurrent components.

ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ FE-SKVIT ДЛЯ КЛАССИФИКАЦИИ ВИДОВ ЦИФРОВОЙ МОДУЛЯЦИИ НА ОСНОВЕ ДАТАСЕТА RADIOML 2018.01A

асп. Ушакова А.М., проф. Приоров А.Л.

Ярославский государственный университет им. П.Г. Демидова

Решается задача автоматического распознавания видов цифровой модуляции радиосигналов с использованием методов глубокого обучения на общедоступном эталонном датасете RadioML 2018.01A. Актуальность темы обусловлена необходимостью создания высокоточных алгоритмов для систем когнитивного радио и мониторинга радиоэфира. На вход сети подавались I/Q-компоненты в исходном виде после преобразования в тензор размерности (2, 1024), а на ее выходе формировался вектор длиной 5, соответствующий пяти классам популярных видов цифровой модуляции: BPSK, QPSK, 16PSK, 16QAM и 64QAM. Активация Softmax применялась для получения апостериорных вероятностей принадлежности входного сигнала к каждому из указанных классов.

Введение

Современное развитие систем беспроводной связи характеризуется усложнением сигнально-помеховой обстановки и необходимостью гибкого использования достаточно ограниченного радиочастотного спектра. В этой связи задачи когнитивного радио, предполагающие автоматическое распознавание типов сигналов, становятся все более актуальными. Классификация видов цифровой модуляции является ключевым этапом для обработки сигнала в условиях априорной неопределенности. Сравнительный анализ различных нейросетевых архитектур для решения данной задачи, в том числе в условиях медленных замираний сигнала, проведен авторами ранее [1, 2].

Традиционные методы распознавания модуляции, основанные на анализе статистических моментов и спектральных характеристик, часто требуют экспертного подбора признаков и показывают недостаточную эффективность в условиях низкого отношения сигнал/шум (ОСШ).

С развитием методов глубокого обучения, особенно сверточных нейронных сетей, появилась возможность автоматического извлечения наиболее информативных признаков непосредственно из сырых данных [3]. Одной из стандартных баз данных радиосигналов для обучения и тестирования нейросетевых методов является открытый датасет RadioML 2018.01A, содержащий большой объем размеченных данных с различными видами модуляции и широким диапазоном ОСШ [4, 5]. В данной работе исследуется применение достаточно новой нейросетевой архитектуры FE-SKVIT для решения задачи классификации на этом датасете.

Целью исследования является оценка эффективности нейросетевой архитектуры FE-SKVIT [6]. В ходе работы обучена и протестирована модель FE-SKVIT на 5 видах модуляции из датасета RadioML 2018, при различных отношениях сигнал/шум (ОСШ) от -20 до +30 дБ с шагом 2 дБ.

Обучение и тестирование нейросетевого алгоритма

Как уже отмечалось, для обучения и тестирования нейросетевой модели использовался общедоступный датасет RadioML 2018.01A. Исходный датасет содержит 24 класса модуляции (19 цифровых и 5 аналоговых) в диапазоне ОСШ от –20 до +30 дБ, однако в настоящей работе использован более ограниченный набор данных. Выбраны 5 базовых цифровых модуляций (BPSK, QPSK, 16PSK, 16QAM, 64QAM), широко применяемых с современных системах и сетях беспроводной связи. Датасет разделен на обучающую (80%), валидационную (10%) и тестовую (10%) выборки с сохранением баланса классов.

Общее количество сигналов составило 80 000 (по 16 000 на каждый вид модуляции), что существенно меньше полного датасета. Это ограничение связано с вычислительными ресурсами используемого для обучения компьютера: процессор Intel Core i5-8250U с базовой частотой 1,60 ГГц (турбо до 1,80 ГГц), 8 ГБ оперативной памяти, 64-разрядная архитектура x64. Обучение осуществлялось на центральном процессоре.

Модель обучалась с использованием оптимизатора Adam и функции потерь категориальной кросс-энтропии. Для предотвращения переобучения применялась регуляризация Dropout и ранняя остановка обучения.

Для классификации сигналов выбрана архитектура FE-SKVIT (Feature Extraction - SKVIT). Она адаптирована для работы с двумерным представлением сигнала (I/Q компоненты как два канала размером 2x1024). Ключевой особенностью модели является комбинация сверточных слоев для извлечения локальных временных признаков и механизма внимания типа Squeeze-and-Excitation (SE), который позволяет адаптивно перекалибровать канальные признаки, фокусируясь на наиболее информативных из них.

Результаты обучения и тестирования нейросетевого алгоритма

В ходе обучения фиксировались потери и точность как на обучающей, так и на тестовой выборке. Динамика обучения представлена на рисунках 1 и 2. Как видно, модель демонстрирует устойчивую сходимость: потери на обучающей и тестовой выборках монотонно убывают, а точность растёт, достигая стабильного уровня к 35-й эпохе. Отсутствие значительного расхождения между обучающей и тестовой выборками указывает на отсутствие переобучения.

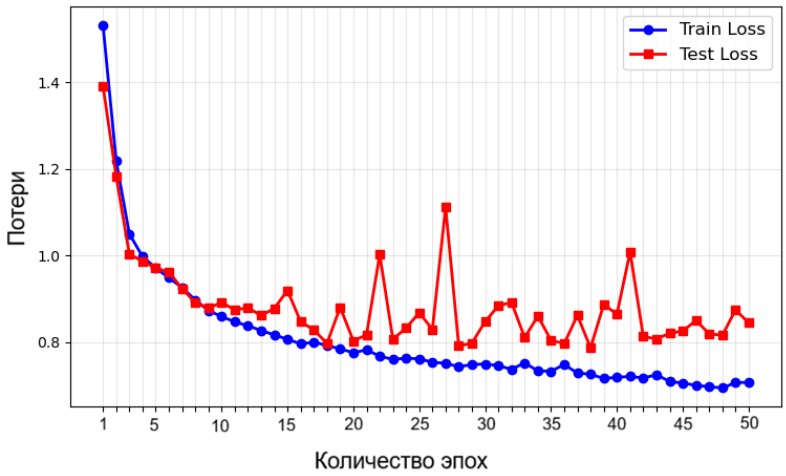


Рисунок 1 – График зависимости потерь от количества эпох

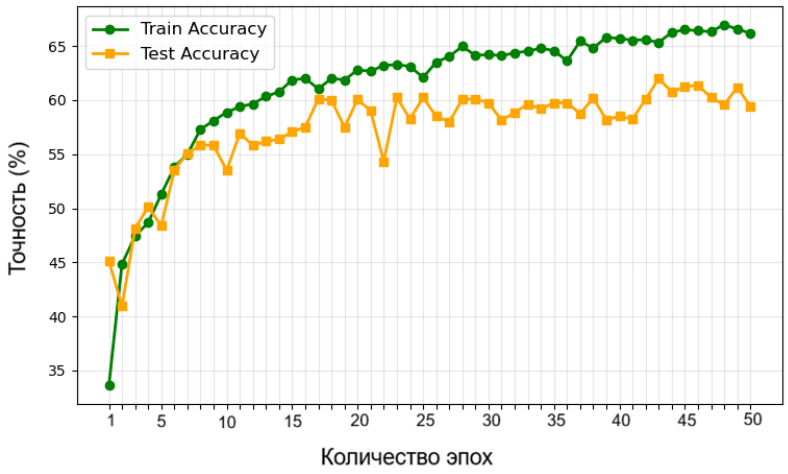


Рисунок 2 – График зависимости точности распознавания вида модуляции от количества эпох

Для более детального анализа сходимости построен комплексный график, включающий четыре стандартных графика: динамику потерь, динамику точности, а также разницу между обучающей и тестовой выборками по потерям и точности. Соответствующие данные представлены на рисунке 3.

Заполненные зелёным и красным цветами области на нижних графиках визуализируют зоны переобучения и недообучения. В данном случае разница остаётся небольшой и колеблется около нуля, что подтверждает устойчивость обученной модели.

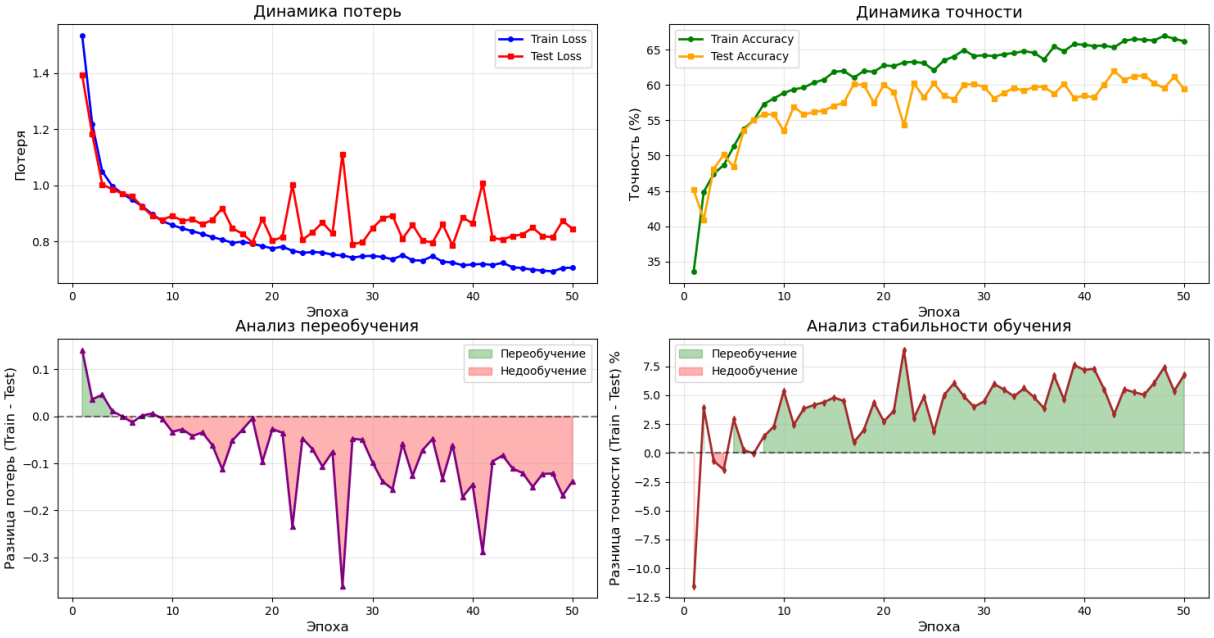


Рисунок 3 – Анализ сходимости обучения модели FE-SKVIT

На рисунке 4 представлена нормализованная матрица ошибок, которая демонстрирует, между какими видами модуляции возникают основные ошибки распознавания.

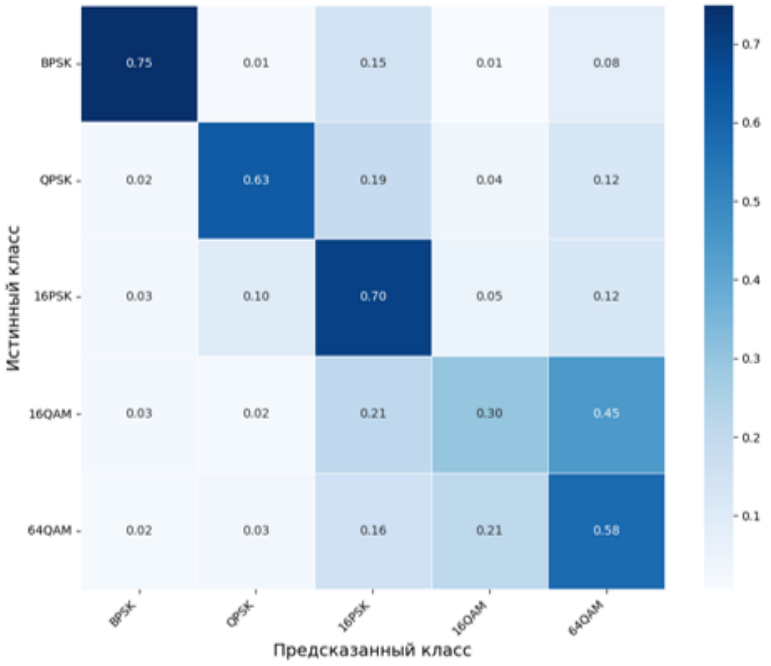


Рисунок 4 – Матрица ошибок нейросетевой модели FE-SKVIT

Наиболее частые ошибки возникают между модуляциями с одинаковой геометрией (круговой для PSK и квадратной для QAM), но разным числом точек, в частности, между 16QAM и 64QAM

(ошибки составляют 45% для 16QAM и 21% для 64QAM) и между 16PSK и QPSK (ошибки составляют 10% для 16PSK и 19% для QPSK).

Заключение

Обучена и протестирована нейросетевая архитектура FE-SKVIT с использованием датасета RadioML 2018.01A. Полученные результаты подтверждают высокую эффективность предложенного подхода: модель продемонстрировала устойчивую классификацию пяти распространенных видов цифровой модуляции в широком диапазоне отношений сигнал/шум. Максимальная точность распознавания вида модуляции на обучающей выборке составила 66,20%. На тестовой выборке – 59,44%.

Достигнутые показатели точности позволяют рассматривать данную архитектуру в качестве перспективного решения для систем когнитивного радио, программно-определяемого радио и систем радиомониторинга.

Направления дальнейших исследований включают расширение номенклатуры распознаваемых видов модуляции, оптимизацию архитектуры сети, а также анализ поведения модели в условиях замираний сигнала.

Литература

1. Ushakova A., Priorov A., Lavrentev M., Kupriyanov E. A Comparative Analysis of Neural Network Architectures for Automatic Digital Modulation Classification under Fading Conditions // 2025 IEEE XVII International Scientific and Technical Conference on Actual Problems of Electronic Instrument Engineering (APEIE). Novosibirsk. Russian Federation. 2025. pp. 1-6.

2. Приоров А.Л., Ушакова А.М. Тимонькин М.В. Разработка нейросетевого алгоритма распознавания вида модуляции радиосигнала в условиях медленных замираний в радиоканале // Цифровая обработка сигналов. 2025. № 4. С. 11-15.

3. O'Shea T.J., Corgan J., Clancy T.C. Convolutional Radio Modulation Recognition Networks // Engineering Applications of Neural Networks: 17th International Conference, EANN 2016. Aberdeen, UK, 2016. Proceedings. Springer International Publishing, 2016. P. 213-226. – DOI: 10.1007/978-3-319-44188-7_16.

4. RF datasets for machine learning [Электронный ресурс]. Режим доступа: https://www.deepsig.ai/datasets (Дата обращения: 15.01.2026).

5. O'Shea T.J., Roy T., Clancy T.C. RadioML 2018.01A: A Synthetic Dataset for Machine Learning in Wireless Communications // 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Kalamata, Greece, 2018. P. 1-5. – DOI: 10.1109/SPAWC.2018.8445945.

6. Zheng G., Zang B., Yang P., Zhang W., Li B. FE-SKViT: A Feature-Enhanced ViT Model with Skip Attention for Automatic Modulation Recognition // Remote Sensing. 2024. Vol. 16. № 22. Article 4204.

APPLICATION OF THE FE-SKVIT NEURAL NETWORK FOR CLASSIFICATION OF TYPES OF DIGITAL MODULATION BASED ON THE RADIOML 2018.01A DATASET

Ushakova A., Priorov A.

P.G. Demidov Yaroslavl State University

The problem of automatic recognition of types of digital modulation of radio signals is solved using deep learning methods on the publicly available RadioML 2018.01A reference dataset. The relevance of the topic is due to the need to create high-precision algorithms for cognitive radio and broadcast monitoring systems. The input of the network was supplied with I/Q components in their original form after conversion to a tensor of dimension (2, 1024), and at its output a vector of length 5 was formed, corresponding to five classes of popular types of digital modulation: BPSK, QPSK, 16PSK, 16QAM and 64QAM. Softmax activation was used to obtain a posteriori probabilities of the input signal belonging to each of the specified classes.

МЕТОДЫ АНАЛИЗА ТЕКСТОВЫХ ДАННЫХ В ЗАДАЧАХ ПРОФЕССИОНАЛЬНОЙ ОРИЕНТАЦИИ ШКОЛЬНИКОВ

студ. Чистякова П.М.

Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева

В работе рассматриваются методы анализа текстовых данных для решения задачи автоматизированной профессиональной ориентации школьников на основе текстовых описаний навыков и интересов. Для классификации профессиональных направлений исследуются классические алгоритмы машинного обучения, рекуррентные нейронные сети и современные трансформерные модели. В рамках исследования выполнено сравнение подходов на основе TF-IDF и логистической регрессии, двунаправленной рекуррентной сети BiLSTM, а также модели BERT в двух вариантах применения: без дообучения в режиме извлечения признаков и с дообучением на целевой задаче классификации. Экспериментальные исследования проведены на открытом наборе данных, содержащем перечни профессиональных навыков и соответствующие направления деятельности в сфере информационных технологий. Полученные результаты показали преимущество трансформерных моделей, при этом наилучшие значения метрик были достигнуты при дообучении модели BERT, что подтверждает эффективность применения предварительно обученных языковых моделей в задачах профессиональной ориентации.

Введение

В условиях постоянных изменений на рынке труда и разнообразия образовательных программ перед старшеклассниками стоит сложная и ответственная задача – сделать обоснованный выбор направления подготовки при поступлении в университет. Этот выбор во многом определяет дальнейшую профессиональную траекторию и успешное начало карьеры. Традиционные методы профориентации, основанные на тестировании, экспертных консультациях или анкетировании, зачастую оказываются трудоемкими и ограниченными в масштабах анализа.

Развитие методов обработки естественного языка открывает новые возможности для анализа текстовой информации, связанной с образовательными предпочтениями абитуриентов. Современные модели – от классических подходов машинного обучения до архитектур на основе трансформеров – позволяют эффективно извлекать смысловые признаки, тематические структуры и скрытые паттерны в текстах учащихся, что может служить основой для построения персонализированных рекомендаций при выборе направления подготовки.

Обзор литературы

Анализ текстовой информации для выявления образовательных и профессиональных предпочтений относится к области автоматического понимания естественного языка (Natural Language Understanding, NLU) и тематического моделирования. Исторически данная задача решалась с помощью методов, основанных на частотном анализе лексики и ручном выделении признаков, однако их точность была ограничена неспособностью улавливать контекст и семантику. Прорыв в этой сфере связан с переходом от классических алгоритмов машинного обучения к распределённым представлениям и глубоким нейронным сетям, способным извлекать сложные лингвистические зависимости. Современный этап развития характеризуется доминированием архитектур на основе трансформеров и предобученных языковых моделей, которые задают новый стандарт для задач классификации и интерпретации текстов, в том числе в образовательной сфере.

В работе [1] рассматривается система профессиональной ориентации обучающихся, основанная на методах классического машинного обучения. Для формирования рекомендаций используются структурированные данные, полученные с помощью анкетирования, включающего оценку личностных характеристик и учебных предпочтений. Авторы применяют алгоритмы Random Forest, Decision Tree и Support Vector Machine и показывают их применимость для задач карьерного консультирования. Вместе с тем, ориентация преимущественно на анкетные данные ограничивает возможности анализа неструктурированных текстовых описаний интересов и навыков обучающихся.

В работах [2, 3] рассматривается применение больших языковых моделей в задачах автоматизированного консультирования и поддержки принятия решений при выборе направления обучения в высшем образовании. В работе [2] предложена система на основе

LLM, предназначенная для автоматизации типовых обращений абитуриентов и формирования персонализированных рекомендаций по выбору образовательных программ. Для формализации задачи используется сопоставление профиля обучающегося с направлениями подготовки на основе векторных представлений и метрик семантического сходства, а экспериментальная оценка проводится с применением подхода Retrieval-Augmented Generation и языковых моделей LLAMA, GPT и DEEPSEEK. Полученные результаты демонстрируют повышение эффективности и качества консультационной поддержки абитуриентов. В работе [3] исследуется возможность использования LLM для поддержки принятия решений путём сравнительного анализа рекомендаций, сформированных моделью GPT-4, и рекомендаций университетских консультантов. Оценка, выполненная экспертами по критериям полезности, обоснованности и соответствия профилю студента, показала, что рекомендации модели во многих случаях сопоставимы с экспертными и частично с ними совпадают, что подтверждает потенциал применения LLM в задачах академического и профессионального консультирования при сохранении контроля со стороны человека.

Таким образом, анализ существующих работ показывает, что задачи выявления образовательных и профессиональных предпочтений обучающихся решаются с использованием широкого спектра подходов – от классических методов машинного обучения, основанных на структурированных анкетных данных, до современных систем на базе больших языковых моделей, ориентированных на диалоговую поддержку и формирование рекомендаций.

В работе будет рассмотрено практическое применение описанных методов к реальной задаче автоматизированной профессиональной ориентации школьников. Основное внимание уделено этапам подготовки и обработки текстовых данных, построению признаковых представлений с использованием TF-IDF и языковых моделей, настройке и обучению различных классификаторов, а также оценке их эффективности на основе стандартных метрик качества. Кроме того, рассматривается возможность использования предварительно обученных трансформерных моделей с дообучением на целевой задаче.

Методы анализа текстовых данных

Для решения задачи классификации профессиональных направлений на основе текстовых описаний навыков в работе будут использоваться три различных подхода, отличающихся уровнем абстракции представления текста и архитектурой модели машинного обучения: классический статический метод на основе TF-IDF с логистической регрессией, рекуррентная нейронная сеть с двунаправленной архитектурой BiLSTM, а также трансформерная модель BERT, основанная на механизме самовнимания. Использование нескольких методов позволяет оценить влияние модели и способа представления текста на качество классификации.

В рамках классического подхода тексты преобразуются в векторное представление с использованием меры TF-IDF (Term Frequency – Inverse Document Frequency), которая отражает важность термина в конкретном документе относительно всего корпуса текстов. Значение TF-IDF для термина t в документе d вычисляется как произведение частоты термина в документе и логарифма обратной документной частоты:

$$TF - IDF(t,d) = TF(t,d)\log\frac{N}{DF(t)}, \quad (1)$$

где N – общее количество документов в корпусе, а DF(t) – количество документов, содержащих термин t.

Полученное разреженное векторное представление подается на вход классификатора логистической регрессии, который оценивает вероятность принадлежности текста каждому из профессиональных направлений.

Другой подход основан на рекуррентных нейронных сетях, а именно архитектуре двунаправленной долгой краткосрочной памяти (BiLSTM, Bidirectional Long Short-Term Memory). В отличие от TF-IDF, рекуррентные модели обрабатывают текст в виде последовательности токенов. Каждый текст предварительно преобразуется в последовательность индексов слов, которые затем отображаются в плотные векторные представления фиксированной размерности с помощью слоя эмбеддингов. Архитектура

BiLSTM включает два рекуррентных слоя LSTM, обрабатывающих входную последовательность в прямом и обратном направлениях, что позволяет учитывать как предшествующий, так и последующий контекст для каждого слова. Итоговое скрытое представление последовательности подается на полносвязный слой классификации с функцией активации Softmax. Выход модели представляет собой распределение вероятностей принадлежности входного текста каждому из классов профессиональной направленности. Общая структура модели BiLSTM для классификации текстов представлена на рисунке 1.

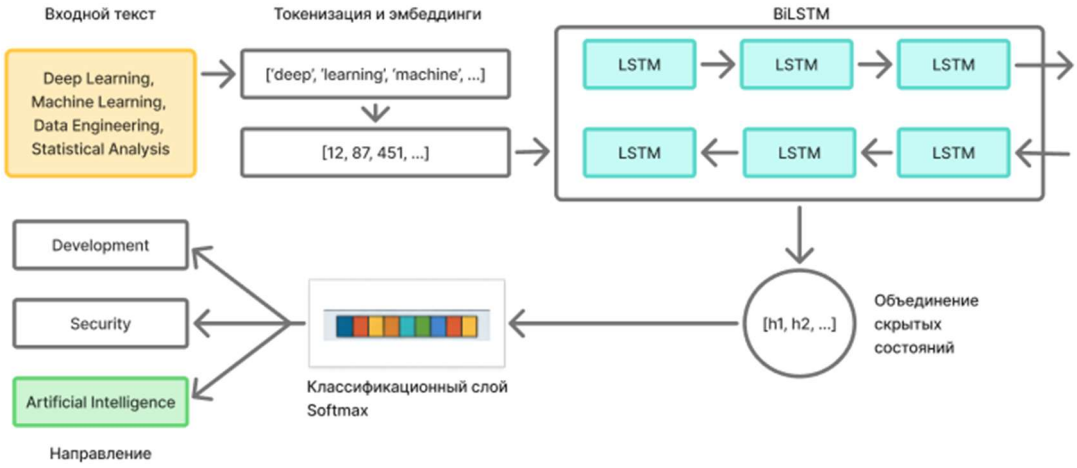


Рисунок 1 – Общая структура модели BiLSTM

Наиболее современным методом анализа текстовых данных в данной работе является трансформерная модель BERT (Bidirectional Encoder Representations from Transformers). В основе данной архитектуры лежит механизм самовнимания, позволяющий модели учитывать взаимосвязи между всеми токенами входной последовательности независимо от их позиции. Текст с перечнем навыков подается на вход модели в виде последовательности токенов, дополненной специальными служебными токенами, после чего обрабатывается стеком трансформерных блоков. Для решения задачи классификации используется скрытое представление специального токена [CLS], которое подается на полносвязный классификационный слой с функцией активации Softmax.

Экспериментальные исследования

Для проведения экспериментальных исследований в работе был использован открытый набор данных, размещенный на платформе Mendeley Data [4], содержащий информацию о профессиональных навыках студентов и выбранных ими направлениях будущей деятельности в сфере информационных технологий. Набор данных был дополнительно модифицирован: англоязычные наименования профессиональных направлений заменены на эквивалентные названия на русском языке, что позволило согласовать используемые классы с формулировками, используемыми в российской системе образования.

В практическом сценарии школьникам предлагается пройти краткий профориентационный тест в формате анкеты, оценивающей интерес к ключевым профессиональным навыкам, таким как управление базами данных, анализ данных, разработка ПО, проектирование интерфейсов, информационная безопасность, машинное обучение. На основе ответов формируется текстовое описание профиля интересов, которое подаётся на вход классификационной модели, результатом работы которой является отнесение школьника к одному или нескольким профилям профессиональной направленности.

Перед проведением экспериментов набор данных был предварительно обработан: тексты приведены к единому формату, удалены повторяющиеся записи, а также выполнена нормализация наименований направлений. В наборе данных представлены шесть классов профессиональной деятельности, в каждом из которых достаточное количество примеров для обучения моделей и корректной оценки качества классификации. Данные были разделен на обучающую и тестовую выборки в соотношении 80% и 20% соответственно с сохранением пропорций классов. Описание набора данных представлено в таблице 1.

Таблица 1. Описание набора данных

Направление профессиональной деятельности	**Количество записей с перечнем навыков**
Разработка программного обеспечения	1148
Искусственный интеллект	804
Программная инженерия	714
Информационная безопасность	580
Анализ данных	558
Проектирование пользовательских интерфейсов	272

Для решения задачи автоматической классификации текстов были рассмотрены несколько подходов, различающихся используемыми методами представления текста и сложностью моделей. В качестве базового метода применялся классический подход на основе TF-IDF-векторизации и логистической регрессии. В качестве нейросетевого решения рассматривалась рекуррентная модель BiLSTM, учитывающая последовательность слов. Основное внимание в работе уделялось трансформерной модели BERT, использующей контекстные эмбеддинги и механизм самовнимания. Были исследованы два варианта её применения: без дообучения – в режиме извлечения признаков с использованием представления токена [CLS] и последующей классификацией логистической регрессией, а также с дообучением на целевой задаче, при котором параметры трансформерных слоёв адаптировались под специфику текстов навыков.

Для обоих вариантов использования BERT тексты навыков предварительно токенизировались с помощью стандартного токенизатора соответствующей модели, при этом максимальная длина входной последовательности была ограничена 128 токенами, что оказалось достаточным для всех записей в наборе данных. В задаче классификации при дообучении BERT использовался дополнительный выходной слой с функцией Softmax, обеспечивающей отнесение каждого текстового описания к одному из шести классов профессиональной деятельности. Дообучение модели проводилось в течение трёх эпох с использованием оптимизатора AdamW и фиксированной скорости обучения. Все вычислительные эксперименты выполнялись в среде Google Colab с использованием графического ускорителя NVIDIA T4 и библиотеки HuggingFace Transformers.

Для оценки качества классификации использовались метрики точности (Accuracy), а также Macro- и Weighted-усреднённые значения F1-меры. Метрика Accuracy отражает долю правильно классифицированных примеров, однако при наличии классов с различным количеством объектов она может не в полной мере характеризовать качество модели. В связи с этим дополнительно применялась Macro F1-мера, обеспечивающая равный вклад каждого класса в итоговую оценку, и Weighted F1-мера, учитывающая дисбаланс классов за счёт взвешивания по числу примеров.

Для одного класса F1-мера определяется как гармоническое среднее между точностью (Precision) и полнотой (Recall). Precision показывает, какая доля объектов, отнесенных моделью к данному классу, действительно относится к нему, а Recall характеризует способность модели находить все объекты, принадлежащие заданному классу. F1-мера вычисляется по формуле:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}, \quad (2)$$

Macro F1-мера вычисляется как среднее арифметическое F1-мер по всем классам:

$$F1_{macro} = \frac{1}{K} \sum_{i=1}^{K} F1_i, \quad (3)$$

где K – количество классов.

Weighted F1-мера вычисляется как взвешенное среднее F1-мер с весами, пропорциональными числу объектов каждого класса:

$$F1_{weighted} = \sum_{i=1}^{K} w_i F1_i, \quad (4)$$

где w_i – отношение количества объектов класса к общему количеству объектов в выборке.

В таблице 2 представлены результаты сравнительной оценки качества классификации профессиональных направлений на основе текстовых описаний навыков с использованием различных моделей машинного обучения и нейросетевых архитектур. Рекуррентная модель BiLSTM продемонстрировала сопоставимые, но несколько более низкие значения метрик по сравнению с классическим методом на основе TF-IDF и логистической регрессии, что может быть обусловлено ограниченным объёмом обучающей выборки. Использование модели BERT без дообучения позволило повысить качество классификации по сравнению с рекуррентной архитектурой, подтвердив эффективность контекстных языковых представлений даже без адаптации модели к конкретной задаче. Наилучшие значения всех рассматриваемых метрик были достигнуты при использовании модели BERT с дообучением, что свидетельствует о преимуществе трансформерных моделей при анализе кратких и средних по длине текстовых описаний навыков.

Таблица 2. Сравнительная оценка качества классификации профессиональных направлений

Модель	**Точность**	**Macro F1-мера**	**Weighted F1-мера**
TF-IDF + Logistic Regression	0,944	0,950	0,940
BiLSTM	0,936	0,940	0,940
BERT (feature-based)	0,940	0,940	0,940
BERT (fine-tuning)	0,957	0,960	0,960

Для более детального анализа качества классификации была построена матрица ошибок для модели BERT с дообучением, представленная на рисунке 2.

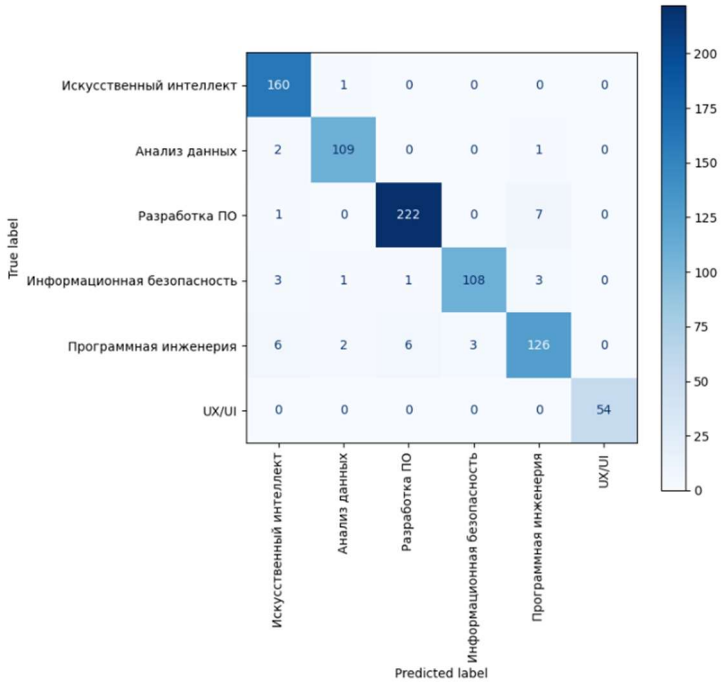


Рисунок 2 – Матрица ошибок для BERT (fine-tuning)

Анализ матрицы показывает, что основная масса объектов корректно классифицируется, о чём свидетельствует преобладание значений на главной диагонали. Минимальное количество ошибок наблюдается для направления «Проектирование пользовательских интерфейсов» (UX/UI), что объясняется высокой специфичностью используемой терминологии. Наибольшее число ошибок возникает между близкими по содержанию направлениями, такими как «Разработка программного обеспечения» и «Программная инженерия», а также между «Искусственный интеллект» и «Анализ данных», что обусловлено пересечением используемых технологий и навыков в текстовых описаниях.

На основании полученных экспериментальных результатов можно сделать вывод о целесообразности применения предварительно обученных трансформерных моделей для задач профессиональной ориентации школьников. Модель BERT с дообучением продемонстрировала наибольшую способность к обобщению и корректному выявлению семантической связи между описываемыми навыками и соответствующими направлениями профессиональной деятельности. В то же время результаты, полученные при использовании BERT без дообучения, а также классических и рекуррентных моделей, показывают, что при ограниченных вычислительных ресурсах и небольшом объёме данных возможны альтернативные подходы, обеспечивающие приемлемый уровень качества классификации.

Заключение

Методы анализа текстовых данных, основанные на современных нейросетевых архитектурах, позволяют эффективно решать задачи профессиональной ориентации школьников. В данной работе было проведено сравнительное исследование классических алгоритмов машинного обучения, рекуррентных нейронных сетей и трансформерных моделей в задаче классификации профессиональных направлений по текстовым описаниям навыков. Эксперименты выполнялись на открытом наборе данных, отражающем профессиональные предпочтения в сфере информационных технологий. Дополнительно набор данных был доработан для совмещения с вопросами и направлениями обучения на русском языке, что позволило использовать его совместно с профориентационными тестами для школьников. В ходе исследования были рассмотрены два варианта применения модели BERT: без дообучения в режиме извлечения контекстных эмбеддингов, и с дообучением на целевой задаче классификации. Результаты показали, что использование трансформерной модели BERT обеспечивает более высокие значения метрик точности и F1-меры по сравнению с классическими и рекуррентными подходами, при этом наилучшие результаты достигаются при дообучении модели на предметно-ориентированном наборе данных. Таким образом, применение предварительно обученных языковых моделей с тонкой настройкой является перспективным направлением для построения интеллектуальных систем поддержки выбора образовательного и профессионального направления при поступлении в университет.

Литература

1. Panthee, S., Rajkarnikar, S., & Begum, R. (2023). Career Guidance System Using Machine Learning. *Journal of Advanced College of Engineering and Management*, *8*(2), 113–119.
2. N. N. H. Van, P. H. Do, V. N. Hoang, T. T. K. Nguyen and M. T. Pham, "AI-Powered University Admission Counseling: A Use Case of Large Language Models in Student Guidance," in *IEEE Transactions on Learning Technologies*, vol. 18, pp. 856-868, 2025, doi: 10.1109/TLT.2025.3604096
3. Lekan, K., & Pardos, Z. A. (2025). AI-Augmented Advising: A Comparative Study of GPT-4 and Advisor-based Major Recommendations. *Journal of Learning Analytics*, *12*(1), 110-128.
4. Hossain Faruque, Sakir; khushbu, sharun akter (2024), “Student Career Dataset”, Mendeley Data, V2, doi: 10.17632/4spj4mbpjr.2

METHODS OF TEXT DATA ANALYSIS IN PROFESSIONAL GUIDANCE TASKS FOR SCHOOLCHILDREN

Chistyakova P.M.

Siberian State University of Science and Technology
named after academician M.F. Reshetnev

The paper considers methods of text data analysis for solving the problem of automated professional guidance for schoolchildren based on textual descriptions of skills and interests. For the classification of professional domains, classical machine learning algorithms, recurrent neural networks, and modern transformer-based models are investigated. The study presents a comparative analysis of approaches based on TF-IDF with logistic regression, a bidirectional recurrent BiLSTM network, and the BERT model used in two application modes: without fine-tuning as a feature extraction model and with fine-tuning for the target classification task. Experimental studies were conducted on an open dataset containing descriptions of professional skills and corresponding career domains in the field of information technologies. The obtained results demonstrate the advantage of transformer-based models, with the best metric values achieved by the fine-tuned BERT model, confirming the effectiveness of pretrained language models for professional guidance tasks.

МЕТОДЫ ОБНАРУЖЕНИЯ ДЕФЕКТОВ МЕТАЛЛИЧЕСКИХ ПОВЕРХНОСТЕЙ С ПРИМЕНЕНИЕМ YOLOV8

студ. Ильгов А.Т.

Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева

В работе рассматривается задача автоматического обнаружения дефектов поверхности металлических изделий, актуальная для систем промышленного контроля качества. Анализируются методы повышения надёжности компьютерного зрения при детекции визуальных дефектов металла. Исследовано влияние добавления негативных примеров из датасета Severstal к обучающему набору NEU-DET на качество детекции дефектов металла с использованием YOLOv8. Показано, что расширение выборки снижает число ложных срабатываний: Precision выросла на 7.5 % при сохранении mAP. Дальнейшая оптимизация параметров аугментации (Mosaic, HSV-Value, Scale, Translate) повысила Precision до 0.691, а также улучшила детекцию сложных классов — rolled-in_scale и crazing. Результаты подтверждают эффективность включения негативных примеров и подбора параметров аугментации для повышения надёжности моделей в промышленном контроле качества.

Введение

Металлические изделия играют ключевую роль во многих отраслях промышленности, включая машиностроение, строительный сектор и металлургическое производство. В условиях ужесточения требований к надежности и долговечности продукции возрастает значимость своевременного выявления дефектов поверхности металла. Такие повреждения, как трещины, царапины и коррозия, способны существенно ухудшать эксплуатационные характеристики изделий и повышать риск возникновения аварийных ситуаций. В связи с этим актуальной является задача создания автоматизированных систем контроля, обеспечивающих оперативное и точное обнаружение дефектов металлических поверхностей.

Современные методы компьютерного зрения, основанные на использовании свёрточных нейронных сетей, позволяют автоматизировать процесс анализа изображений, получаемых как в режиме реального времени, так и из архивных источников данных. Применение таких подходов способствует повышению точности контроля качества и снижению зависимости от субъективных факторов, связанных с визуальной оценкой оператором.

Обзор литературы

В работе [1] предложен метод автоматической классификации дефектов поверхности стальных листов на основе свёрточной нейронной сети, обученной на датасете Severstal [4]. Модель с четырьмя свёрточными и шестью полносвязными слоями показала точность 96% на обучающей и 73% на тестовой выборке при использовании оптимизатора Adam и функции потерь focal Tversky loss.

В исследовании [2] тот же датасет применялся для задач сегментации дефектов с использованием U-Net и модифицированной Deep Residual U-Net, последняя из которых продемонстрировала лучший результат (Dice = 0.731 против 0.543).
В работе [3] рассмотрен подход к детекции дефектов на основе модифицированной RetinaNet с деформируемыми свёртками, модулем CA-BiFPN и функцией потерь IA-BCELoss.

В исследованиях [2] и [3] обучение проводилось на датасете NEU-DET [5]. Датасет NEU-DET включает шесть типов дефектов металлической поверхности, однако не содержит изображений без дефектов. Отсутствие негативных примеров может отрицательно влиять на способность модели корректно различать дефектные и бездефектные участки поверхности, что снижает её обобщающую способность при практическом применении. С учётом изложенного, представляется целесообразным исследование влияния состава обучающего набора данных на качество детекции дефектов.

Экспериментальные исследования

В рамках данного исследования в качестве базовой модели была выбрана архитектура YOLOv8, обучение которой будет происходить на двух вариантах данных:

1. Исходный датасет NEU-DET (6 классов дефектов);
2. Расширенный датасет, включающий все изображения из датасета NEU-DET, а также негативные примеры (изображения без дефектов) из датасета Severstal.

Структура объединенной обучающей выборки приведена на рисунке 1. Валидационная выборка сформирована аналогичным образом (по 60 изображений на класс, включая негативные).

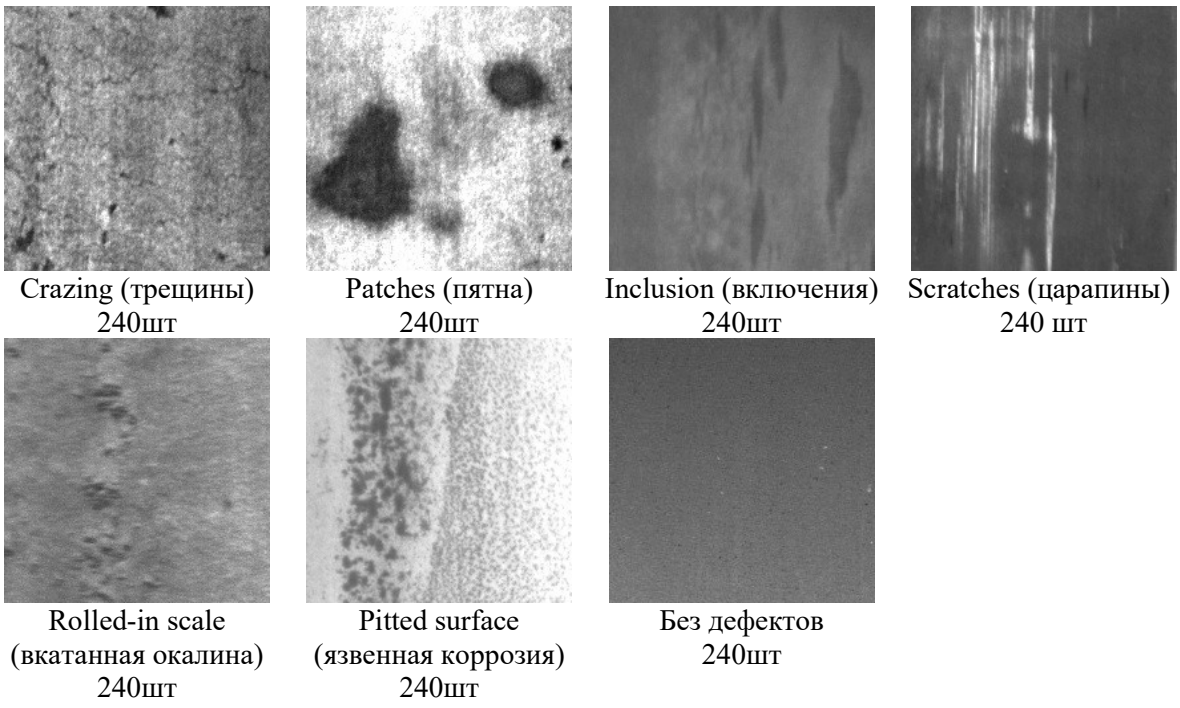


Рисунок 1 – Структура обучающей выборки в объединенном датасете

Эксперименты проводились при фиксированных гиперпараметрах. Значения данных гиперпараметров представлены в таблице 1.

Таблица 1. Фиксированные параметры обучения

Параметр	**Значение**
Imgsz	200
Epochs	100
Batch	16
Optimizer	AdamW
Learning rate	0.001
Seed	42
Patience	0

Оценка качества моделей выполнена по стандартным метрикам:

1. Precision (точность) — доля истинно положительных предсказаний среди всех объектов, классифицированных моделью как дефектные:

$$P = \frac{TP}{TP+FP}, \quad (1)$$

где TP - количество корректно обнаруженных дефектов, FP - количество ложных срабатываний.

2. Recall (полнота) – доля обнаруженных дефектов от общего числа реальных дефектов;

$$R = \frac{TP}{TP+FN}, \quad (2)$$

где FN – количество пропущенных дефектов.

3. mAP50 – среднее значение AP (average precision) по всем классам при IoU ≥ 0.5:

$$\mathrm{IoU} = \frac{\mathrm{Area}(\mathrm{B}_{\mathrm{pred}} \cap \mathrm{B}_{\mathrm{gt}})}{\mathrm{Area}(\mathrm{B}_{\mathrm{pred}} \cup \mathrm{B}_{\mathrm{gt}})}, \quad (3)$$

где $\mathrm{B}_{\mathrm{pred}}$ и B_{gt} – предсказанная и истинная (ground truth) ограничивающие рамки соответственно.

4. mAP50-95 – среднее AP при IoU от 0.5 до 0.95 с шагом 0.05, отражающее устойчивость модели к строгости локализации:

$$IoU_{thr} \in \{0.50, 0.55, 0.60, \ldots, 0.95\}, \quad (4)$$

Полученные результаты приведены таблицах 2 и 3.

Таблица 2. Метрики модели, обученной на исходном датасете NEU-DET

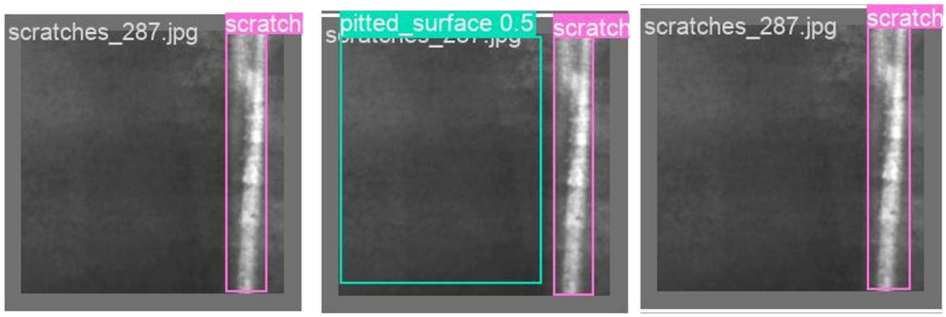
Class	Box (P)	Box (R)	mAP50	mAP50-95
all	0.613	0.704	0.726	0.389
crazing	0.657	0.355	0.531	0.225
inclusion	0.667	0.824	0.837	0.444
patches	0.809	0.943	0.942	0.619
pitted_surface	0.681	0.759	0.797	0.449
rolled-in_scale	0.385	0.437	0.410	0.172
scratches	0.477	0.909	0.837	0.427

Модель демонстрирует высокую полноту (Recall), однако точность (Precision) остается низкой, что свидетельствует о наличии значительного числа ложных срабатываний. Особенно это заметно для классов scratches и patches, где модель склонна генерировать избыточные детекции, не всегда соответствующие реальным дефектам. В то же время классы rolled_in_scale и crazing характеризуются наихудшими значениями mAP50 и mAP50–95, что указывает на слабую способность модели корректно распознавать данные типы дефектов. Значения метрик второй версии модели, обученной на расширенном наборе данных, представлены в таблице 3.

Таблица 3. Метрики модели, обученной на расширенном датасете

Class	Box (P)	Box (R)	mAP50	mAP50-95
all	**0.659**	0.684	0.725	0.389
crazing	**0.738**	0.261	0.500	0.197
inclusion	**0.679**	0.792	0.815	0.454
patches	0.784	0.883	0.923	0.605
pitted_surface	**0.723**	0.701	0.780	0.441
rolled-in_scale	**0.457**	0.530	0.474	0.197
scratches	**0.571**	0.934	0.858	0.440

Метрики mAP50 и mAP50–95 практически не изменились по сравнению с исходной моделью, что говорит о сохранении общей способности к детекции. Однако, общая точность повысилась на 7.5% (с 0.613 до 0.659), что явно указывает на снижение числа ложных детекций и улучшение качества предсказаний. При этом для большинства классов наблюдается более сбалансированное соотношение между Precision и Recall, что свидетельствует об улучшении обобщающей способности модели и её большей надёжности в реальных условиях. Существенное улучшение точности особенно заметно для таких проблемных классов, как crazing и scratches, хотя в ряде случаев это достигнуто за счёт снижения полноты – у crazing Recall упал с 0.355 до 0.261. Такой компромисс может быть интерпретирован как следствие более строгого отбора предсказаний, когда модель предпочитает пропустить часть истинных дефектов, чтобы избежать ошибочных срабатываний. Пример визуального улучшения качества детекции после обучения на расширенном датасете представлен на рисунке 2.



а б в

Рисунок 2 – Пример повышения точности детекции:
а) размеченное изображение из валидационной выборки,
б) результат модели 1 (обученной на датасете без негативных примеров)
в) результат модели 2 (обученной на датасете с негативными примерами)

Далее было проведено исследование влияния параметров встроенной аугментации, реализованной в архитектуре YOLOv8.

В YOLOv8 предусмотрена гибкая система аугментации, включающая множество настраиваемых параметров. Каждый из параметров управляет определённым типом трансформации входных изображений. Параметр **Mosaic** объединяет четыре случайных изображения из обучающего набора в одно составное, что способствует лучшему обучению модели на мелких объектах и повышает её устойчивость к изменению контекста. Параметры **HSV-Value**, **HSV-Hue** и **HSV-Saturation** регулируют яркость, оттенок и насыщенность изображения соответственно, что полезно при работе с цветными изображениями, однако избыточно при работе с изображениями в оттенках серого, поэтому компоненты, связанные с цветовым пространством HSV, за исключением яркости (параметр **Value**), были отключены. Параметры **Scale** и **Translate** контролируют случайное масштабирование и смещение изображения, что должно помочь итоговой модели адаптироваться к возможным вариациям расстояния до поверхности или положения изделия на конвейере.

В рамках эксперимента были выделены четыре параметра аугментации, наиболее значимые для данной предметной области. Для каждого параметра был определен диапазон и шаг изменения:

- **Mosaic:** значение изменялось от 0 (отключено) до 1 (применяется ко всем батчам) с шагом 0.25;
- **HSV-Value:** диапазон от 0 (без изменения яркости) до 0.8 (сильные колебания яркости) с шагом 0.2;
- **Scale:** изменение масштаба от 0 (без масштабирования) до 0.5 (до ±50% изменения размера) с шагом 0.1;
- **Translate:** смещение изображения по горизонтали и вертикали от 0 до 0.1 (до 10% от размера изображения) с шагом 0.02.

Для каждого значения рассматриваемых параметров проводилось независимое обучение модели с нуля при фиксированных остальных параметрах, равных значениям по умолчанию. Также, в экспериментах были фиксированы гиперпараметры, перечисленные в таблице 1.

После завершения каждого цикла обучения модель оценивалась на валидационной выборке по тем же метрикам, что и ранее: Precision, Recall, mAP50 и mAP50–95. Полученные результаты позволили построить зависимости метрик качества от значений каждого из исследуемых параметров. Эти зависимости визуализированы на графиках, представленных на рисунке 3, где по оси абсцисс отложены значения параметра аугментации, а по оси ординат — соответствующие значения метрик.

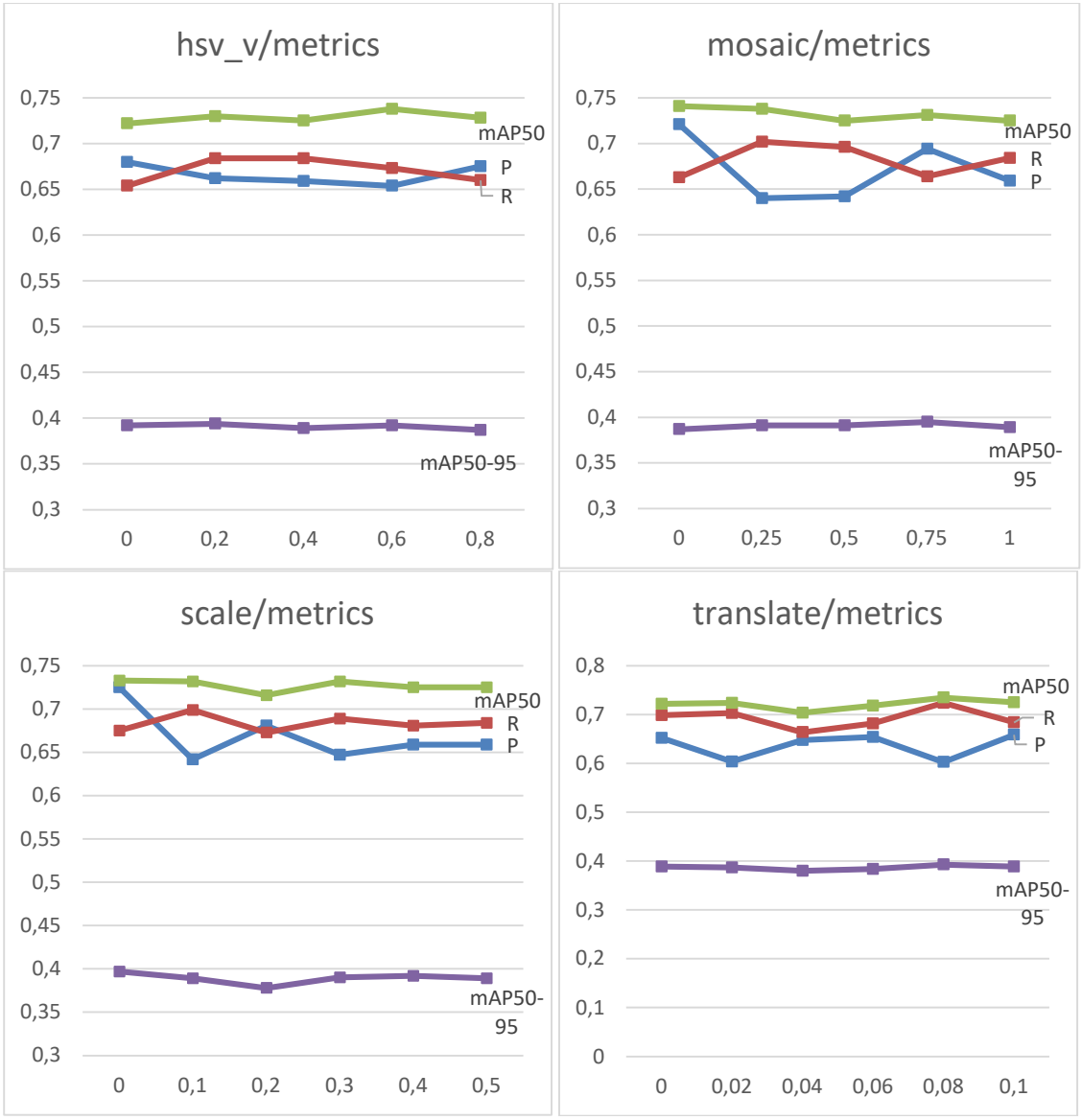


Рисунок 3 – Зависимость метрик итоговой модели от параметров аугментации

На основе анализа полученных зависимостей была сформирована конфигурация параметров аугментации, которые, в рамках проведённых экспериментов, продемонстрировали более сбалансированные результаты по сравнению с настройками по умолчанию: HSV-Value = 0.2, Mosaic = 0.75, Scale = 0, Translate = 0.08, HSV-Hue = HSV-Saturation = 0.0.

С использованием этой конфигурации была обучена финальная версия модели. Её метрики представлены в таблице 4.

Таблица 4. Метрики итоговой модели с заданными параметрами аугментации

Class	Box (P)	Box (R)	mAP50	mAP50-95
all	**0.691**	**0.686**	**0.733**	**0.393**
crazing	0.618	**0.407**	0.485	0.191
inclusion	**0.743**	0.727	0.816	0.466
patches	**0.88**	0.837	0.916	0.603
pitted_surface	0.722	**0.77**	0.783	0.427
rolled-in_scale	**0.59**	0.5	0.562	0.213
scratches	**0.594**	0.876	0.839	0.458

Анализ результатов показывает, что общая метрика mAP50–95 изменилась незначительно по сравнению с предыдущими версиями (увеличение всего на 0.004), однако наблюдаются важные качественные улучшения на уровне отдельных классов. Так, для класса rolled-in_scale, который ранее демонстрировал крайне низкие значения точности (0.385), произошло последовательное и устойчивое её увеличение (0.457 при расширении датасета, 0.590 в финальной модели). Аналогично, для класса crazing удалось достичь более сбалансированного соотношения между Precision и Recall.

Заключение

В работе исследовано влияние расширения обучающей выборки на качество детекции поверхностных дефектов металлических изделий с использованием архитектуры YOLOv8. Показано, что дополнение исходного датасета NEU-DET изображениями без дефектов из датасета Severstal позволяет повысить точность предсказаний (Box Precision) на 7.5 % при практически неизменной полноте и стабильных значениях метрик mAP50 и mAP50–95. Это свидетельствует о значительном снижении числа ложных срабатываний без ущерба для способности модели обнаруживать реальные дефекты. Проведённый анализ параметров аугментации показал, что настройка таких параметров, как Mosaic, HSV-Value и Translate, позволяет добиться улучшения баланса между точностью и полнотой. Предложенная эмпирически обоснованная конфигурация аугментации обеспечила дополнительный рост Precision до 0.691 и улучшение метрик по наиболее проблемным классам, таким как *rolled-in_scale* и *crazing*. Таким образом, применение расширенной обучающей выборки и настройка параметров аугментации демонстрируют практическую эффективность. Эти подходы не требуют усложнения архитектуры модели или сбора новых размеченных данных, но при этом повышают надёжность и применимость системы в реальных производственных условиях.

Литература

1. Aji Teguh Prihatno, Ida Bagus Krishna Yoga Utama, Jun Yong Kim*, and Yeong Min Jang "Metal Defect Classification Using Deep Learning", in IEEE Twelfth International Conference on Ubiquitous and Future Networks, pp. 389-393, 2021, doi: 10.1109/ICUFN49451.2021.9528702
2. Didarul Amin, Shamim Akhter "Deep Learning-Based Defect Detection System in Steel Sheet Surfaces", in IEEE Region 10 Symposium, 2020, doi: 10.1109/TENSYMP50017.2020.9230863
3. Zhanglin Yang, Yu Liu "A steel surface defect detection method based on improved RetinaNet", in Scientific Reports, doi: 10.1038/s41598-025-88527-x
4. Dataset "Severstal: Steel Defect Detection", URL: https://www.kaggle.com/c/severstal-steel-defect-detection (Дата обращения: 15.12.2025)
5. Dataset "NEU Surface Defect Database", URL: https://www.kaggle.com/datasets/kaustubhdikshit/neu-surface-defect-database (Дата обращения: 18.12.2025)

METHODS FOR DETECTING DEFECTS IN METAL SURFACES USING YOLOV8

Ilgov A.T.

Siberian State University of Science and Technology
named after academician M.F. Reshetnev

The paper focuses on the problem of automated detection of defects on metal surfaces. It studies the effect of adding negative samples from the Severstal dataset to the NEU-DET training set on the quality of metal defect detection using YOLOv8. It is shown that expanding the sample reduces the number of false positives: Precision increased by 7.5% while maintaining the mAP. Further optimization of augmentation parameters (Mosaic, HSV-Value, Scale, Translate) increased Precision to 0.691, and improved detection of complex classes — rolled-in_scale and crazing. The results confirm the effectiveness of including negative examples and selecting augmentation parameters to increase the reliability of models in industrial quality control.

МЕТОДЫ СОЗДАНИЯ 3D-МОДЕЛЕЙ В СФЕРЕ ОБРАЗОВАНИЯ

студ. Степанова М.Р.

Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева

В работе проводится анализ технологических особенностей способов построения 3D-моделей, сравнительная оценка преимуществ и недостатков, а также рассматривается их практическая применимость с двух ключевых точек зрения: педагогов, для которых важен критерий простоты интеграции в учебный план, и обучающихся, для которых наиболее значимы доступность и интерактивность модели.

Введение

Обучение – это трудоемкий процесс, который нередко становится повинностью для человека. Традиционные методы передачи знаний зачастую не способны эффективно и наглядно передавать сложные пространственные, динамические и абстрактные концепции, что снижает вовлеченность и глубину понимания. На современном этапе развития цифровых технологий возникает объективная необходимость в новом формате обучения, особенно для поколения, с детства погруженного в цифровую среду.

Одним из ключевых инструментов такой трансформации являются трехмерные модели и технологии виртуальной, дополненной и смешанной реальности, также известные как VR, AR и MR. Они позволяют вывести образовательный процесс за рамки плоских изображений и текста, предоставляя возможность интерактивного исследования, деконструкции и взаимодействия с объектами изучения. Это особенно критично в областях, где обучение сопряжено с высокими рисками, стоимостью ошибки или этическими ограничениями. Например, в медицинском образовании 3D-модели и симуляторы уже доказали свою эффективность, позволяя будущим хирургам отрабатывать сложные операции, а студентам – изучать анатомию в мельчайших деталях без вреда для пациента [1].

Однако потенциал 3D-моделирования гораздо шире. С его помощью можно безопасно моделировать опасные химические и физические эксперименты, например цепные ядерные реакции, проводить тренировки для сотрудников МЧС в виртуальных копиях горящих зданий, обучать ветеринаров на точных цифровых двойниках животных или погружать школьников в исторические реконструкции. Целью работы является систематизация и обзор современных методов создания 3D-моделей, пригодных для использования в образовательном процессе.

Обзор литературы

Для интеграции 3D-контента в образование необходимо понимать спектр доступных технологий его создания. Условно их можно разделить на несколько крупных категорий, каждая из которых обладает уникальным балансом между трудоемкостью, стоимостью, точностью и доступностью.

Анализ современных исследований и практик показывает растущий тренд на геймификацию и визуализацию образования. Исследования в области когнитивной психологии подтверждают, что мультимодальное обучение с использованием интерактивных 3D-объектов способствует лучшему усвоению и долгосрочному запоминанию информации по сравнению с пассивными методами. Разрыв между поколением Альфа и устаревшими образовательными методиками подчеркивает необходимость адаптации учебных программ [2].

В работе [3] в качестве метода построения 3D-модели предложена трёхмерная реконструкция на основе искусственной нейронной сети, использующаяся в качестве пассивного метода, в связи с её относительно низкой стоимостью, универсальностью применения по сравнению с активными методами, требующими специализированного и дорогостоящего оборудования, а также с тем, что она позволяет восстанавливать форму объекта, используя в качестве входных данных только его 2D-проекции или разреженный набор 3D-точек, что актуально для задач компьютерного зрения, медицинской визуализации и работы с архивными материалами**.**

В работе [4] в качестве метода построения 3D-модели предложена трёхмерная реконструкция на основе искусственной нейронной сети. Методами реализации выступили: нейросетевая

сегментация на основе адаптированного алгоритма Deep Watershed Transform для выделения объектов, каскадная обработка с помощью специализированных свёрточных сетей для извлечения геометрических признаков и аналитико-геометрический модуль для построения 3D-моделей с использованием эталонных масштабов. Такой набор методов обеспечивает универсальность подхода и его применимость для восстановления моделей по одиночным 2D-изображениям.

В работе [5] предложено применять ИНС для моделирования памятников архитектуры и зданий, для чего используются современные генеративные модели и нейронные сети для классификации изображений. Ключевыми методами реализации в данной работе являются: двухэтапная стратегия применения нейросетевых инструментов без дообучения и с дообучением, а также специфические технологические конвейеры для каждой проектной задачи.

Обзор методов создания 3D-моделей

Потенциал 3D-моделирования простирается практически на все дисциплины и уровни образования. В медицине и биологии оно позволяет создавать детализированные анатомические атласы, симуляторы для отработки хирургических вмешательств и модели молекул или клеток. В рамках естественных наук, таких как физика и химия, трехмерная визуализация делает наглядными невидимые процессы, например: течение жидкости или силовые поля, и предоставляет возможность безопасно проводить виртуальные эксперименты с опасными веществами или в экстремальных условиях. Для инженерии и архитектуры – это инструмент изучения сложных механизмов, проектирования и инспекции виртуальных макетов сооружений. В гуманитарной сфере, в частности в истории и археологии, технологии позволяют проводить цифровые реконструкции древних городов, артефактов и значимых событий. Наконец, в сфере профессиональной подготовки, например: МЧС, авиация, промышленность, 3D-модели служат основой для высокоточных тренажеров, предназначенных для отработки действий в аварийных и нештатных ситуациях, что значительно повышает безопасность и эффективность обучения.

В зависимости от исходных данных, требуемой точности и доступных ресурсов, создание 3D-моделей для образовательных целей может осуществляться несколькими принципиально разными способами. Условно их можно разделить на три крупные категории: реконструкция из реальных данных, профессиональное моделирование и генерация с помощью искусственного интеллекта.

В сфере цифрового образования методы создания трехмерного контента образуют спектр, простирающийся от точной оцифровки реального мира до генерации объектов силами искусственного интеллекта. Выбор конкретного подхода диктуется задачами, ресурсами и целевой аудиторией.

Реконструкция на основе реальных объектов, таких как фотограмметрия и SfM представляет собой технологию создания точной цифровой копии физического объекта путем обработки множества его фотографий, сделанных с разных ракурсов. Её суть заключается в алгоритме Structure from Motion (SfM), который анализирует перекрывающиеся двумерные снимки, находит общие точки, вычисляет положение камеры для каждого кадра и в итоге восстанавливает трехмерное облако точек. Для реализации этого подхода используются как бесплатные инструменты с открытым исходным кодом, например VisualSFM (VSFM), так и коммерческие решения, такие как Agisoft Metashape или RealityCapture. Бесплатные инструменты, хотя и требуют высоких вычислительных мощностей и определенных технических навыков для работы через консольный интерфейс, предоставляют доступ к мощному функционалу. Коммерческие же программы предлагают полный, удобный графический конвейер обработки, автоматизируя весь процесс от выравнивания снимков до наложения текстур, однако они часто требуют приобретения лицензии. В образовательном контексте эти методы незаменимы для оцифровки реальных учебных экспонатов – от исторических артефактов и геологических образцов до архитектурных памятников для виртуальных экскурсий, позволяя студентам интерактивно изучать объекты, недоступные в физической лаборатории или музее. Конкретный пример реализации SfM-алгоритма можно рассмотреть на примере работы с VisualSFM в работе [6].

Схема классификации методов создания 3D-моделей представлена на рисунке 1.

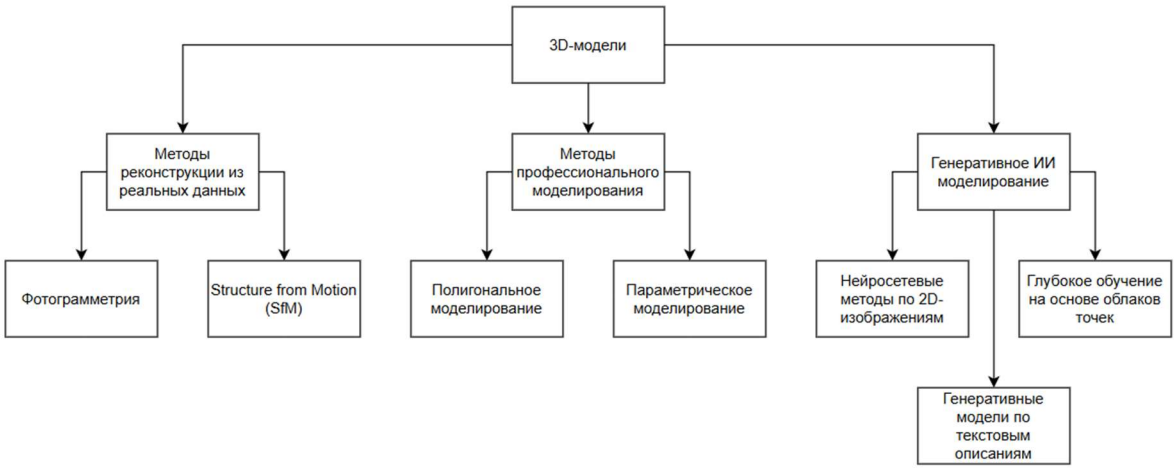


Рисунок 1 – Схема классификации методов создания 3D-моделей

Профессиональное 3D-моделирование и CAD – это методы целенаправленного создания модели с нуля с помощью специализированного программного обеспечения, требующие от пользователя высоких профессиональных навыков. Это направление делится на две основные ветви: полигональное моделирование и CAD-моделирование. Полигональное моделирование в таких программах, как Blender, Maya или 3ds Max, подразумевает ручное создание и редактирование художником вершин, ребер и граней (полигонов) для формирования объекта любой сложности – от гиперреалистичного до стилизованного. Его ключевое преимущество – творческий контроль и возможность моделировать объекты, не существующие в реальности, например для исторических реконструкций или визуализации научных гипотез. С другой стороны, CAD-моделирование, представленное системами вроде Fusion 360, SolidWorks или TinkerCAD, основано на создании параметрических моделей из точных эскизов и операций, например выдавливание и вычитание. Такие модели легко редактируются изменением исходных параметров, что обеспечивает высочайшую точность и делает данный подход идеальным для инженерных и технических дисциплин. В образовании эти методы применяются для создания идеализированных учебных моделей, например идеальной клетки или механизма, обучения инженерному проектированию, а также как самостоятельная учебная дисциплина – компьютерная графика и анимация.

Нейросетевые подходы, такие как генеративный AI, знаменуют собой наиболее быстро развивающееся направление, где трехмерная модель генерируется искусственным интеллектом на основе текстового описания или двухмерного изображения. Этот подход кардинально отличается от предыдущих тем, что пользователь задает не геометрические параметры, а семантическое намерение – цель или концепцию будущего объекта. Облачные сервисы, такие как Meshy.ai, 3DAiStudio или Masterpiece Studio, предоставляют простой веб-интерфейс, в котором пользователь вводит текстовый запрос на подобие «детализированная 3D-модель сердца человека с кровеносными сосудами, учебный стиль» или загружает картинку для преобразования в 3D. Всего за несколько минут сервис выдает готовую текстурированную модель.

Главные достоинства этого метода – крайне низкий порог входа, не требующий специальных навыков моделирования, и очень высокая скорость генерации, что делает его идеальным для быстрого прототипирования. Однако он имеет и существенные ограничения: пользовательский контроль над топологией и деталями модели минимален, результат может быть непредсказуемым или содержать артефакты, а качество пока часто уступает ручной работе.

Важно понимать, что «волшебство» подобных сервисов основано на сложных исследовательских архитектурах нейронных сетей, разработанных в академической среде. Современные модели, лежащие в основе генерации 3D, часто используют принципы диффузионных моделей, которые учатся постепенно преобразовывать случайный шум в связную 3D-структуру, следуя текстовой подсказке. Более того, для работы с мешами – наиболее распространенным и удобным форматом 3D-объектов – создаются специализированные

нейронные сети, способные не только генерировать, но и интеллектуально обрабатывать сложную, нерегулярную структуру полигональных сеток.

Ярким примером такой сети, рассмотренной в работе [7], является DGNet (Dual Graph Network). Её ключевая инновация – пирамида двойственных графов (dual graph pyramid), которая преобразует произвольный, даже дефектный меш, в иерархическую структуру, понятную для нейросети. Это позволяет преодолеть ограничение многих ранних архитектур, работавших только с идеальными сетками.

Сердцем DGNet является гибридная свертка (hybrid convolution), математическая суть которой выражена в формуле объединения признаков для каждой грани f_i:

$$y_i = H_{self}(x_i) + H_{geo}(x_i) + H_{euc}(x_i), \quad (1)$$

где H_{self} – собственные признаки, H_{geo} – геодезические признаки, H_{euc} – евклидовы признаки.

Эта формула описывает, как сеть работает в процессе анализа 3D-объекта:

1. Линейное преобразование признаков самой грани:

$$H_{self}(x_i) = W_1 x_i, \quad (2)$$

где W_1 – обучаемые веса сети.

2. Используется операция максимума для обеспечения инвариантности к перестановкам:

$$H_{ge}\ (x_i) = \max_{j \in N_s(i)} (W_0 x_j) \quad (3)$$

3. Анализ пространственного контекста. Окружающее пространство разбивается на воксели. Для вокселя v, содержащего центр грани f_i, и его 6 соседей вычисляются признаки, которые затем взвешенно суммируются. Это позволяет сети находить связи между изолированными компонентами, расположенными близко в пространстве.

$$H_{euc}(x_i) = \sum_{k=0}^{6} W_k^v x_k^v \quad (4)$$

Такая гибридная архитектура свертки позволяет сети одновременно понимать локальную топологию поверхности и глобальные пространственные отношения, что критически важно для работы с реальными, неидеальными данными.

Архитектура DGNet [7] адаптируется под задачу: для классификации объектов используется нисходящая структура на основе Hybrid Block, а для детальной сегментации – U-Net-подобная архитектура с остаточными Hybrid ResBlock, представленная на рисунке 2.

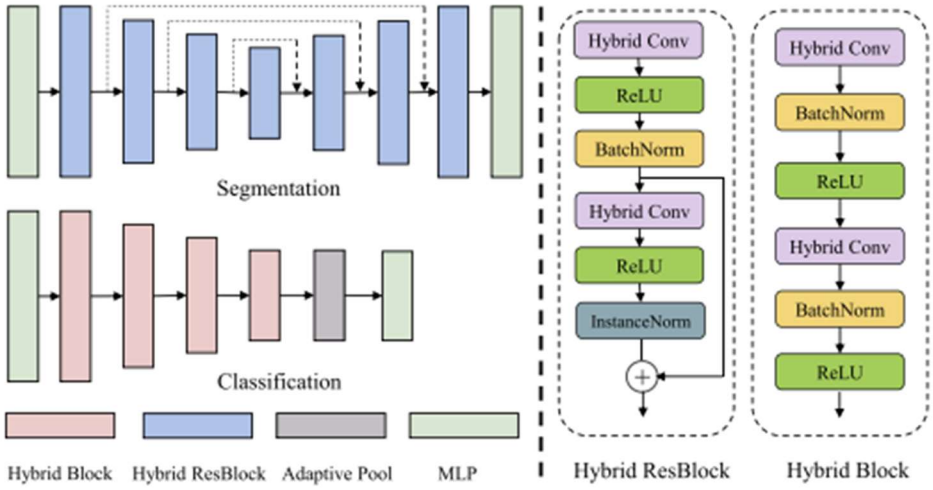


Рисунок 2 – Архитектуры сети классификации и сегментации DGNet

Таким образом, простота генерации по запросу для конечного пользователя скрывает под собой сложные математические аппараты и передовые исследования в области геометрического глубокого обучения. Для образования, особенно для преподавателей-предметников, этот подход открывает уникальные возможности. Он позволяет быстро создавать наглядные пособия,

иллюстрируя сложные понятия по запросу аудитории. Понимание лежащих в его основе принципов и формул, подобных реализованным в DGNet, позволяет не просто использовать AI как черный ящик, но и критически оценивать его результаты, прогнозировать развитие инструментов и осмысленно интегрировать их в исследовательский и учебный процесс, например, для автоматизированного анализа или сегментации сложных 3D-моделей, полученных другими методами.

Рассмотренные методы представляют собой три различных философии создания цифрового контента, каждая из которых занимает свою нишу в образовательном процессе. Реконструкция соединяет физический и цифровой мир, обеспечивая максимальную точность для работы с реальными объектами, но требует физического доступа к образцу и сложной обработки данных. Профессиональное моделирование – это инструмент абсолютного контроля, позволяющий создавать концептуальные и инженерные модели с нуля, но требующий значительных временных затрат и экспертных навыков. Нейросетевые подходы – это технологии семантической абстракции и автоматизации, снижающие барьер входа и позволяющие мгновенно создать необходимое, за счет детализации и полного контроля над результатом.

Эти методы не столько конкурируют, сколько дополняют друг друга в системе образования. Так, нейросеть вроде DGNet может быть использована для автоматической обработки и сегментации мешей, полученных методом фотограмметрии. CAD-модель может послужить высококачественной основой для обучения генеративных AI-моделей. В перспективе именно синергия классических и нейросетевых подходов откроет новые возможности: учитель сможет быстро сгенерировать AI-прототип модели, доработать его в CAD-системе для точности, а затем использовать 3D-сканер для наложения текстур с реального объекта, создавая тем самым комплексные, интерактивные и максимально информативные учебные материалы нового поколения.

Каждый из рассмотренных методов занимает свою нишу в образовательной экосистеме. Фотограмметрия служит мостом между физическим и цифровым миром, профессиональное моделирование является инструментом точного конструирования и глубокого изучения, а генеративный ИИ выступает как катализатор оперативной визуализации идей. Их комбинированное и осмысленное применение открывает новые горизонты для создания интерактивной, наглядной и вовлекающей образовательной среды.

Таблица 1 – Сравнение методов 3D-моделирования объектов

Критерий	Фотограмметрия / SfM	Профессиональное моделирование	Нейросетевые AI-сервисы
Необходимое оборудование	Требуется несколько камер, с разрешением от 20 Мпикс, или система с определением расстояния до объекта	Компьютер, специализированное ПО, например: Blender	Только компьютер с доступом в интернет
Требуемые навыки	Средние	Очень высокие	Минимальные
Скорость создания	Средняя/Высокая	Низкая	Очень высокая
Точность и детализация	Фотореалистичная текстура, точная геометрия реального объекта	Контролируемая (может быть как схематичной, так и гиперреалистичной)	Переменная, часто стилизованная, возможны артефакты
Гибкость	Низкая	Абсолютная	Высокая, но контроль ограничен
Оптимальная область применения	Оцифровка реальных экспонатов, создание виртуальных музеев и экскурсий	Создание эталонных учебных моделей, инженерное и дизайнерское образование	Быстрое создание иллюстративных материалов для лекций и презентаций по запросу

Заключение

Проведенный анализ методов создания 3D-моделей для образования – реконструкции, профессионального моделирования и AI-генерации – с позиций педагога и обучающегося показывает, что синергия этих подходов необходима. Для педагога ключевым критерием является простота интеграции, где нейросетевые сервисы, благодаря семантическому интерфейсу и скорости, открывают возможности для оперативного создания контента, в то время как фотограмметрия и CAD-моделирование остаются незаменимы для оцифровки и точного конструирования. Для обучающегося главное – доступность и интерактивность, которые обеспечиваются всеми методами через форматы VR/AR, причем AI позволяет персонализировать модели под запрос. Стратегическая интеграция методов – быстрая AI-генерация прототипа, его точная доработка в CAD и обогащение деталями через оцифровку формирует перспективный конвейер для создания комплексных, педагогически эффективных материалов. Таким образом, осознанное комбинирование технологий 3D-моделирования становится основой для цифровой трансформации образования, позволяя максимально раскрыть потенциал обучения в подготовке специалистов будущего.

Литература

1. Хидиров Г.Т. Результаты обучения с использованием моделей и анимации для улучшения понимания студентов структуры и функции белков / Вестник педагогического университета (Естественные науки) // Хидиров Г.Т, Бобизода Г.М., Баротзода К.А., №3 (15), с. 157-166, 2022

2. Данилова Л.Н. Образовательный запрос поколения альфа / Известия Саратовского университета. Новая серия. Серия: Акмеология образования. Психология развития. // Данилова Л.Н., №1 (45), с. 58-67, 2023

3. Peng L., Shamsuddin S., 3D Object Reconstruction and Representation Using Neural Networks the Association for Computing Machinery, pp. 139-147, 2004

4. Гвоздев О.Г. Нейросетевой метод построения трехмерных моделей ригидных объектов по спутниковым изображениям / Мехатроника, автоматизация, управление // Гвоздев О.Г., Козуб В.А., Кошелева Н.В., Мурынин А.Б., Рихтер А.А., №22(1), с. 48-55, 2021

5. Максимова С.В. Тестирование искусственных нейронных сетей для задач в области архитектуры / Информационные технологии и архитектура // Максимова С.В., Семина А.Е., Райзих Е.В., Белякова Д.П., №4(73), с. 294-307, 2025

6. Даминов И. А. Комбинирование алгоритмов SFM и ORB при 3D-реконструкции / Электронные библиотеки // Даминов И. А., Арсенюк А. Ю., Тощев А. С., №26(4), с. 456–465, 2023

7. Li X., Liu Z., Chen T., Mu T., Ralph R. Martin, Hu S. Mesh Neural Networks Based on Dual Graph Pyramids, pp. 4211 – 4224, 2023

METHODS OF CREATING 3D-MODELS IN THE FIELD OF EDUCATION

Stepanova M.R.

Siberian State University of Science and Technology
named after academician M.F. Reshetnev

The study analyzes their technological features, provides a comparative assessment of their advantages and disadvantages, and examines their practical applicability from two key perspectives: that of educators, for whom the ease of integration into the curriculum is an important criterion, and that of learners, for whom the accessibility and interactivity of the model are most significant.

ДЕТЕКТИРОВАНИЕ ШУМОВ СЕРДЦА С ПОМОЩЬЮ АЛГОРИТМА НА ОСНОВЕ НЕЙРОННОЙ СЕТИ BEATS

доц. Топников А.И.

Ярославский государственный университет им. П.Г. Демидова

Рассматривается применение нейронной сети BEATs, базирующейся на архитектуре трансформерного кодировщика, в задаче детектирования шумов при аускультации сердца. Дообучение и тестирование нейронной сети осуществляется с использованием общедоступной части базы сигналов CirCor DigiScope. Особое внимание уделяется методике тестирования алгоритма, направленной на повышение достоверности результатов и возможности их сравнения с другими. Процедура дообучения, состоящая в коррекции весов всех слоев нормализации и обучении заключительного классификационного слоя, затрагивает малую долю весов модели. Достигнутое значение взвешенной точности классификации пациентов (0,832) позволяет сделать вывод о перспективности развиваемого подхода.

Введение

Современный этап развития методов выявления шумов сердца, основанных на классификации звуковых сигналов аускультации, во многом связан с появлением относительно крупного размеченного набора звуковых сигналов CirCor DigiScope [1, 2], а также с проведением в 2022 году на его основе международного научного конкурса [3]. В дальнейшем уровень взвешенной точности классификации, достигнутый лидерами этого конкурса, был превзойден рядом исследователей. Многие успешные решения основывались на применении предобученных нейронных сетей, построенных по архитектуре Трансформер [4-7]. Данная работа продолжает это направление исследованием перспективности применения алгоритма на основе нейронной сети BEATs (от англ. Bidirectional Encoder representation from Audio Transformers) [8], а также затрагивает вопросы оценки и сопоставления результатов, полученных с использованием общедоступной части базы сигналов CirCor DigiScope.

База сигналов и протокол тестирования

Наиболее крупной размеченной базой сигналов для задачи детектирования шумов сердца является база CirCor DigiScope, записанная с использованием цифрового стетоскопа Littmann 3200. Полная ее версия содержит 5282 звуковые записи аускультации сердца 1568 пациентов [2]. Суммарная длительность записей превышает 33 часа. Однако в общем доступе находится лишь часть базы, содержащая 3163 звуковые записи аускультации сердца 963 пациентов [1]. Остальная часть базы использовалась в слепом тестировании алгоритмов участников научного конкурса 2022 года и не была опубликована после завершения конкурса. Это усложняет процесс разработки новых методов и существенно затрудняет сопоставление результатов, полученных разными исследователями.

Для большинства пациентов в базе представлено несколько звуковых сигналов длительностью от 5 до 168 секунд [2]. Запись сигналов преимущественно производилась в 4 традиционных точках аускультации сердца, соответствующих митральному, аортальному, пульмональному и трехстворчатому клапанам. Для некоторых пациентов приводится не полный набор записей. Также возможно наличие двух записей одного пациента, сделанных в одной точке аускультации, а также сигналов, записанных в иных точках. Частота дискретизации сигналов составляет 4 000 отсчетов в секунду. Каждая запись снабжена файлом с разметкой фаз сердечного цикла и файлом с общей информацией о записи. Также приводится файл формата csv, содержащий обезличенную информацию о всех пациентах: как антропометрическую, так и специализированную, включая данные о наличии или отсутствии шумов сердца, их характере и точках аускультации, в которых они фиксируются и наиболее выражены.

Научный конкурс 2022 года включал в себя две задачи: детектирование шумов и предсказание клинического исхода [3]. Первая задача, связанная с классификации зашумленных звуковых сигналов, может рассматриваться в качестве основной. Именно она в настоящий момент

привлекает основное внимание исследователей, так как кроме очевидной практической значимости связана с решением важных научных задач обработки звуковых сигналов, включая работу с несбалансированными и сильно зашумленными данными. Это исследование также посвящено решению первой из двух задач.

В качестве ключевой метрики для сравнения алгоритмов детектирования шумов сердца в конкурсе 2022 года обоснованно предложена взвешенная точность классификации пациентов со следующими весами: 1 – для класса «Absent» (отсутствие шумов), 5 – для класса «Present» (наличие шумов), 3 – для класса «Unknown» (соответствует случаям, когда врачи, осуществляющие разметку базы, не могли однозначно определить наличие или отсутствие шумов). Особо стоит отметить, что метрика рассчитывается на основе решений, принятых для пациентов, а не для отдельных записей или их фрагментов. Использование в ряде опубликованных исследований только метрик, вычисленных для отдельных сигналов или фрагментов, а также для выборок, существенно отличающихся от исходной базы по соотношению классов, не позволяет оценить достигнутый уровень и сравнить его с результатами других исследователей.

Кроме того, относительно малый объем базы сигналов, ее несбалансированность по классам, а также использование весов при вычислении ключевой метрики приводят к тому, что даже незначительные отличия в матрицах ошибок могут приводить к ощутимым изменениям взвешенной точности. В итоге многократные повторения обучения одной и той же нейронной сети могут давать существенную вариацию значений ключевой метрики. В работе [5] предложен протокол тестирования, подразумевающий 5-кратное повторение процесса обучения для каждого из 3 вариантов хорошо стратифицированных разбиений общедоступной части базы сигналов CirCor DigiScope. Итоговые значения метрик рассчитываются как усреднение метрик по всем 15 запускам. Применение этого подхода приводит к многократному росту временных затрат, что может рассматриваться как плата за повышение достоверности значений метрик и обеспечение лучшей сопоставимости результатов. Также стоит учитывать, что он предъявляет к алгоритмам более строгие требования. Вместо максимизации метрик для отдельных «удачных» запусков обучения нейронной сети, требуется обеспечить стабильно высокое качество обучения каждого запуска. Это требование оказалось трудновыполнимым при использовании подхода на основе кодировщика сети Whisper, представленного в [6, 7], и стимулировало поиск иных решений.

Нейронная сеть BEATs

Нейронные сети, построенные на основе трансформерной архитектуры, нашли широкое распространение не только в задачах распознавания речи и идентификации диктора, но и в задаче классификации звуковых сигналов. Примером может служить нейронная сеть BEATs, демонстрирующая высокие значения метрик для целого ряда наборов звуковых данных [8].

Один из важнейших и наиболее сложных аспектов, возникающих при применении архитектуры Трансформер для обработки звуковых сигналов, – механизм формирования последовательности токенов. В отличие от задачи распознавания речи, где каждый токен, как правило, соответствует определенному временному участку сигнала и содержит в себе информацию, извлеченную из всего доступного диапазона частот, в задаче классификации звуковых сигналов, токены часто соответствуют некоторой прямоугольной области (патчу) мел-спектрограммы, то есть определенному интервалу как по времени, так и по частоте. Такой подход наиболее близок к методам обработки изображений, следствием чего является частое использование вариаций трансформерной архитектуры, созданных для распознавания изображений (например, ViT и SWIN), при создании нейронных сетей для классификации звуковых сигналов. Однако вопросы обучения таких сетей для звуковых данных имеют свою специфику, определяемую природой сигналов.

Для обучения сети BEATs на базе звуковых сигналов AudioSet используется оригинальный итеративный подход, основанный на формировании псевдометок, а не на реконструкции сигнала. На старте токенизация осуществляется случайно инициализированным проекционными

токенизаторами, производится обучение трансформерного кодировщика. Затем происходит обучение токенизатора. Такое чередование обучений токенизатора и кодировщика повторяется несколько раз, пока наблюдается существенное улучшение качества классификации.

Структура нейронной сети построена на основе ViT-кодировщика. Входной звуковой сигнал с частотой дискретизации 16 000 отсчетов в секунду преобразуется в 128-канальную логарифмированную мел-спектрограмму, вычисленную с использованием окна длиной 25 мс и шагом 10 мс. Для преобразования мел-спектрограммы в последовательность токенов она разбивается на патчи (блоки) размера 16×16, которые линейно преобразуются в векторы и нормализуются (рис. 1). Последовательность векторов преобразуется кодировщиком, состоящим из 12 трансформерных блоков, и поступает на полносвязный слой классификации с функцией softmax на выходе. Модель содержит около 90 миллионов параметров.



Рисунок 1 – Схема нейронной сети BEATs

Обученная на базе сигналов AudioSet сеть BEATs затем может дообучаться для целевого набора сигналов за счет обучения заключительного классификационного слоя. Такой подход существенно менее требователен к параметрам вычислительных средств и объему данных.

Дообучение и тестирование нейронной сети

Относительно небольшой объем базы CirCor DigiScope затрудняет обучение большой нейросетевой модели. Поэтому целесообразно использовать схему, в которой предобученный кодировщик (в данном случае – BEATs) выступает в качестве блока извлечения признаков звуковых сигналов. Наиболее простым вариантом является дообучение классификационного слоя при фиксированных («замороженных») коэффициентах кодировщика. Однако результаты предыдущих исследований показывают, что дообучение слоев нормализации в кодировщике, построенном по архитектуре Трансформер, параллельно с обучением классификационного слоя позволяет добиться существенного улучшения итоговой точности классификации [7]. Этот подход использовался и в данном исследовании.

Перед стартом процесса обучения все звуковые сигналы разбиваются на отрезки длиной 5 секунд с шагом 2,5 секунды. Частота дискретизации повышается до 16 000 отсчетов в секунду. Далее сигналы подаются на вход предобученной нейронной сети BEATs (в версии $\text{BEATs}_{\text{iter3+}}$). Для ее дообучения используется оптимизатор, построенный на базе метода AdamW [9] с коэффициентом, отвечающим за скорость обучения, равным 0,001. Размер батча равен 16. Длительность каждого запуска процесса обучения составляет 20 эпох. Величины смещений классификационного слоя инициализируются с учетом доли каждого класса в обучающей выборке. В качестве функции потерь используется категориальная кросс-энтропия. Исследуется три конфигурации: без весов классов, с весами классов, определяемыми их значимостью (те же веса, что используются при вычислении взвешенной точности по методике научного конкурса 2022 года [3]), и с весами классов, определяемыми частотой этих классов (аналогично [5]).

Обработка, включая вычисление потерь, производится для 5-секундных фрагментов. Для вычисления значений метрик на уровне исходных сигналов выходы классификатора до операции softmax (так называемые логиты), соответствующие всем фрагментам одного сигнала, усредняются и затем к ним применяется операция softmax.

Исследуется вариант обучения без аугментации данных, а также с использованием метода mixup [10] с параметром α, равным 0,2; 0,4 и 0,8. Для каждого запуска обучения сохраняется модель, соответствующая эпохе, для которой взвешенная точность классификации сигналов валидационного множества является максимальной. Затем значения метрик, соответствующие наилучшим эпохам, усредняются по всем 15 запускам для каждой конфигурации процесса обучения (таблица 1). Выбор конфигурации для итогового тестирования осуществляется на основе достигнутой на валидационном множестве взвешенной точности классификации сигналов. В случае равенства этих значений для разных конфигураций решение принимается по дополнительной метрике – полноте (recall), усредненной по классам.

Таблица 1. Значения взвешенной точности / полноты на валидационном множестве

Конфигурация процесса обучения		**Без весов**	**Веса определяются значимостью классов**	**Веса определяются частотой классов**
Без аугментации		0,828 / 0,607	0,838 / 0,627	0,830 / 0,678
С аугментацией mixup	α=0,2	0,834 / 0,614	**0,839** / 0,652	0,833 / 0,668
	α=0,4	0,830 / 0,610	0,839 / 0,641	0,828 / 0,664
	α=0,8	0,831 / 0,609	0,838 / 0,638	0,824 / 0,675

На этапе тестирования значения метрик вычисляются как для отдельных записей, так и для пациентов. Принятие решения по пациенту производится на основе простого правила: если хотя бы на одной из записей детектированы шумы сердца, то пациент относится к классу «Present». Далее аналогичное правило применяется для класса «Unknown». Таким образом пациент относится к классу «Absent», если все его записи отнесены классификатором к этому классу. Значения метрик, рассчитанных на уровне пациентов, приводится в таблице 2. Выбор метрик, отображаемых в таблице, обусловлен необходимостью проведения сравнения с другими алгоритмами. Полный набор метрик, вычисляемых в рамках исследования, гораздо шире. Методика тестирования предложенного решения в основных моментах соответствует методике тестирования алгоритма на основе нейронной сети M2D [5]. Методика тестирования алгоритма, основанного на нейронной сети wav2vec 2.0 [4], отличается от описанного выше подхода в сторону упрощения. Значение метрик для этих двух методов взяты из публикаций [4] и [5].

Таблица 2. Результаты тестирования алгоритма и его сравнение с другими

Алгоритм	**Взвешенная точность**	**Полнота (recall)**			
		Среднее значение	**По классам**		
			Absent	**Present**	**Unknown**
На основе BEATs	0,832	0,676	0,851	0,954	0,224
На основе M2D	0,832	0,713	0,868	0,911	0,361
На основе wav2vec 2.0	0,80	0,70	0,83	0,86	0,41

Значения полноты по классам и их среднего стоит оценивать с учетом несбалансированности данных: преобладания пациентов без шумов сердца. Предложенный метод стабильно плохо справляется с детектированием наименее представленного класса «Unknown». Именно этим объясняется низкое среднее значение полноты. Решение проблемы видится в подборе методов аугментации, функции потерь и весов классов, используемых при ее вычислении, а также использовании ансамбля классификаторов. В целом, текущие результаты показывают перспективность используемого подхода и работ по дальнейшему совершенствованию представленного нейросетевого алгоритма и методики его обучения.

Заключение

Таким образом, проведенные исследования демонстрируют перспективность использования нейронной сети BEATs в задаче выявления шумов сердца. Дообучение небольшой доли

коэффициентов нейронной сети позволило достичь взвешенной точности классификации пациентов, равной 0,832. Дальнейшие исследования будут направлены на совершенствование предложенного подхода, включая повышение точности классификации наименее представленного из трех классов.

Литература

1. Oliveira J., Renna F., Costa P., Nogueira M., Oliveira A.C., Elola A., Ferreira C., Jorge A., Bahrami Rad A., Reyna M., Sameni R., Clifford G., Coimbra M. The CirCor DigiScope phonocardiogram dataset // PhisioNet [Электронный ресурс]. Режим доступа: https://physionet.org/content/circor-heart-sound/1.0.3/
2. Oliveira J.H., Renna F., Costa P.D., Nogueira M., Oliveira C., Ferreira C., Jorge A., Mattos S., Hatem T., Tavares T., Elola A., Rad A.-B., Sameni R., Clifford G., Coimbra M. The CirCor DigiScope dataset: from murmur detection to murmur classification // IEEE Journal of Biomedical and Health Informatics. V. 26. N. 6. 2021. P. 2524-2535.
3. Reyna M., Kiarashi Y., Elola A., Oliveira J., Renna F., Gu A., Perez Alday E.A., Sadr N., Mattos S., Coimbra M., Sameni R., Bahrami Rad A., Koscova Z., Clifford G. Heart murmur detection from phonocardiogram recordings: The George B. Moody PhysioNet Challenge 2022 // PLOS Digital Health. V. 2. N. 9. 2023. 22 p.
4. Panah D.S., Hines A., McKeever S. Exploring wav2vec 2.0 model for heart murmur detection // 31st European Signal Processing Conference (EUSIPCO). P. 1010-1014.
5. Niizumi D., Takeuchi D., Ohishi Y., Harada N., Kashino K. Exploring pre-trained general-purpose audio representations for heart murmur detection // 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2024. 4 p.
6. Топников А.И., Смирнова Т.А. Дообучение кодировщика нейронной сети Whisper с помощью промт-векторов и блоков адаптации для задачи обнаружения шумов сердца // Цифровая обработка сигналов. 2025. №2. С. 69-74
7. Топников А.И., Смирнова Т.А. Применение кодировщика нейронной сети Whisper для детектирования шумов сердца // Электроника в медицине. Мониторинг, диагностика, терапия (ЭМ-25): сборник научных статей по материалам XIV Международного симпозиума. Курск, 2025. С. 28-32.
8. Chen S., Wu Y., Wang C., Liu S., Tompkins D., Chen Z., Wei F. BEATs: audio pre-training with acoustic tokenizers // International Conference on Machine Learning (ICML). 2023. 16 p.
9. Defazio A., Yang X., Khaled A., Mishchenko K., Mehta H., Cutkosky A. The road less scheduled // Advances in Neural Information Processing Systems, 37. 2024. P. 9974-10007
10. Zhang H., Cisse M., Dauphin Y.N., Lopez-Paz D. Mixup: beyond empirical risk minimization // International Conference on Learning Representations. 2018. 13 p.

USING THE BEATS NEURAL NETWORK FOR HEART MURMUR DETECTION

Topnikov A.I.

Yaroslavl State University named after P.G. Demidov

The paper investigates the use of the BEATs transformer-based neural network for heart murmur detection from phonocardiogram recordings. Fine-tuning and evaluation are performed on the public subset of the CirCor DigiScope database. A key focus is the testing methodology designed to ensure result reliability and comparability. The fine-tuning procedure, which involves adjusting the weights of all layer normalization layers and training the final classification layer, affects only a few percent of the model's weights. The achieved per patient weighted accuracy (0.832) demonstrates the promise of the developed approach.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРИМЕНЕНИЯ НЕЙРОСЕТЕВЫХ АРХИТЕКТУР В ЗАДАЧЕ ДЕТЕКЦИИ ОБЛАСТИ ИНТЕРЕСА В КОЛОНОСКОПИЧЕСКИХ ВИДЕОДАННЫХ

асп. Котов Н.В.

Ярославский государственный университет им. П.Г. Демидова

Рассматривается задача контроля качества колоноскопических исследований на основе нейросетевого детектирования области купола слепой кишки. Для ее решения исследуются современные детекторы на основе сверточных нейронных сетей и трансформеров – YOLOv11, RTMDet, TOOD и DAB-DETR. Компьютерное моделирование проводилось на собственном наборе данных, содержащем 6927 пар изображений и файлов разметки, а также изображений фона. По результатам тестирования наилучшие результаты продемонстрировала архитектура YOLOv11 по метрикам mAP@50, mAP@50-95, F1-мере показав результаты 0,994, 0,755 и 0,993 соответственно. Исследование послужит основой для разработки модуля автоматического контроля полноты выполнения колоноскопических исследований в реальных клинических условиях.

Введение

В области медицинской визуализации и анализа видеоэндоскопических данных одной из ключевых задач является автоматизация контроля качества диагностических процедур [1]. В частности, при проведении колоноскопических исследований толстой кишки критически важным является подтверждение достижения эндоскопом купола слепой кишки (КСК), что служит индикатором подтверждения полноты осмотра. Визуальное сходство этой анатомической области с другими складками кишечника делает задачу её автоматической детекции нетривиальной [2-3]. Современные методы глубокого машинного обучения, в частности детекторы объектов на основе сверточных нейронных сетей (СНС) и архитектур типа «трансформер», демонстрируют высокую эффективность в подобных задачах анализа видеоданных. Целью данной работы является сравнительное исследование и адаптация передовых архитектур детекторов для надежного обнаружения КСК на колоноскопических видеоизображениях. Адаптация в данной работе заключалась в подборе специфичных параметров обучения и техник аугментации данных, а также в обучении на экспертно размеченном наборе данных с учётом визуального сходства целевой области с фоновыми структурами. Реализация данного подхода позволит создать основу для системы автоматизированного контроля качества проведения соответствующей процедуры.

Используемые архитектуры нейронных сетей

В исследовании использовались четыре современные архитектуры, представляющие различные подходы к задаче детекции объектов.

YOLOv11 – эволюция известного семейства одноэтапных детекторов YOLO, сочетающая скоростную эффективность с улучшенной точностью за счет оптимизированной архитектуры базовой сети (backbone) и промежуточной сети (neck), а также продвинутых техник аугментации и подхода, не использующего явные якорные рамки (anchor-free) в последних версиях [4].

RTMDet – представляет собой семейство высокоэффективных детекторов объектов реального времени, разработанное на основе эмпирического анализа ключевых компонентов, включая архитектуру модели, стратегии назначения меток, аугментации данных и оптимизацию. Архитектурной основой RTMDet служит модифицированный CSPDarkNet, в котором базовый блок усилен depth-wise свертками с большим ядром (5×5) для увеличения рецептивного поля и улучшения контекстного моделирования, что повышает точность без значительного роста вычислительной сложности. Для балансировки скорости и точности, в RTMDet сокращено количество блоков в backbone и расширена ширина модели, а также обеспечена сбалансированная пропускная способность между backbone и neck (CSP-PAFPN), что улучшает распределение вычислений [5-6].

TOOD (Task-aligned One-stage Object Detection) – это одностадийный детектор объектов, разработанный для решения проблемы пространственного рассогласования между задачами классификации и локализации, которые в традиционных детекторах обучаются параллельными

независимыми ветвями [7].

DAB-DETR (Dynamic Anchor Box DETR) – представитель семейства DETR (Detection Transformer), который преодолевает проблемы медленной сходимости оригинальной архитектуры DETR за счет использования динамически обновляемых якорных рамок (anchor boxes) в качестве входных запросов декодера. Эта архитектура сочетает преимущества трансформеров для моделирования глобального контекста с явной позиционной привязкой, характерной для классических детекторов [8].

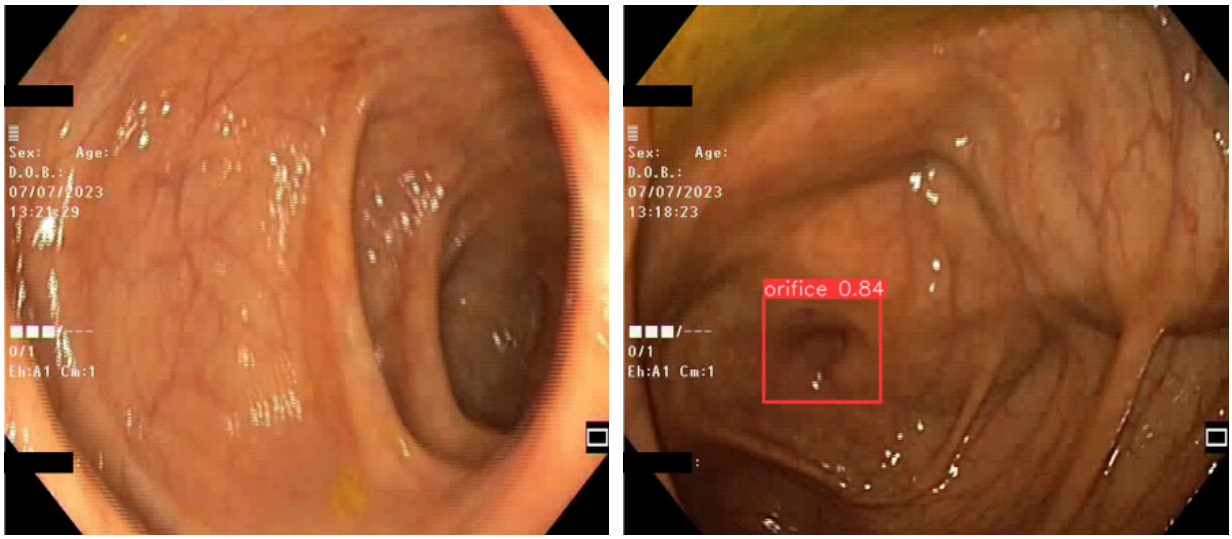
Описание набора данных

В данном исследовании сформирована база изображений, содержащих купол слепой кишки на основе экспертно размеченных видеоданных колоноскопических исследований, проведенных в эндоскопическом отделении ГБУЗ «Ярославская областная клиническая онкологическая больница». Использовались видеопотоки с эндоскопических систем OLYMPUS и PENTAX. Набор данных содержит 4618 пар изображений с размерами 740x576 пикселей и файлов разметки с целевым объектом купола слепой кишки. Также использовались 2309 изображений фона без целевого объекта, но со сложными анатомическими участками, что необходимо для снижения риска ложных срабатываний. В таблице 1 представлена статистика по распределению и количеству объектов в наборе данных.

Таблица 1. Статистика набора данных для обучения и тестирования алгоритмов детектирования

Подмножество	Всего объектов	КСК	«Фон»	Распределение, %
Обучающее	5541	3665	1876	80
Валидационное	692	477	215	10
Тестовое	694	476	218	10
Итог	6927	4618	2309	100

На рисунке 1 изображены примеры кадров из набора данных, содержащие фоновую область и купол слепой кишки.



а) б)

Рисунок 1 – Примеры кадров из эндоскопического исследования с: а) фоном б) предсказанным детектором КСК

Для увеличения объема обучаемых данных и улучшения обобщающей способности моделей применялись различные техники аугментации, специфичные для каждой архитектуры, но имеющие общие черты:

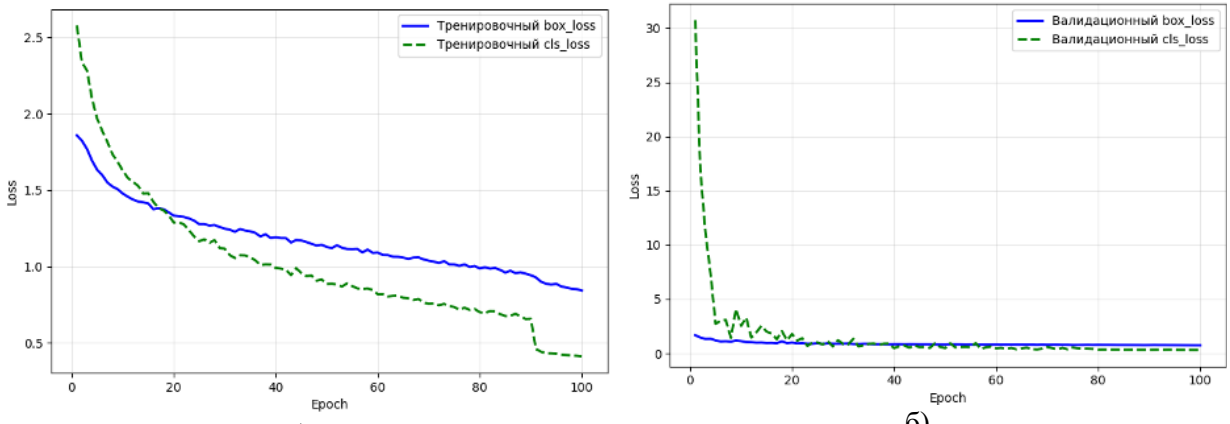
- пространственные преобразования – случайные повороты (до ±10°), сдвиги (до 10%), масштабирование (от 0,4 до 2,0), горизонтальное отражение с вероятностью 0,5;
- цветовые искажения – коррекция в пространстве HSV (оттенок, насыщенность, яркость);
- составные аугментации – Mosaic (комбинирование нескольких изображений в одно) и MixUp (линейная интерполяция пар изображений и их меток) с заданной вероятностью;
- специфичные техники - RandomAffine, PhotoMetricDistortion, RandomCrop.

Результаты исследования

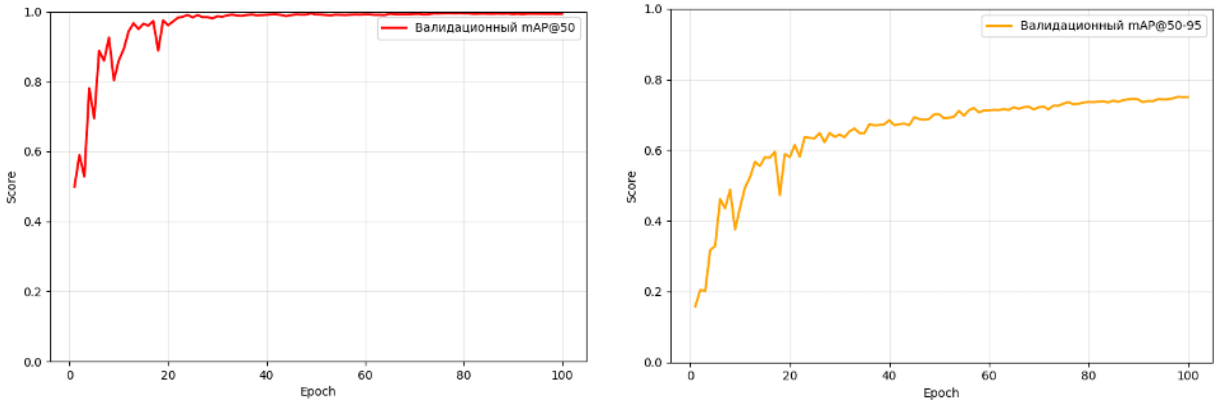
Обучение и тестирование нейронных сетей проводилось с использованием суперкомпьютера NVIDIA DGX-1 на одной видеокарте NVIDIA Tesla V100 с 16 Гб видеопамяти. Все архитектуры обучались в течение 100 эпох с размером батча 16, за исключением DAB-DETR, где размер батча равнялся 4 из-за требований к видеопамяти. В качестве оптимизатора преимущественно использовался AdamW [9] с начальной скоростью обучения 0,001, за исключением TOOD, где применялся алгоритм SGD. Для регуляризации применялся способ L2-регуляризация весов и обрезка градиентов. Обучение контролировалось по функциям потерь, специфичным для каждой архитектуры – QualityFocalLoss + GIoULoss для RTMDet, комбинация FocalLoss [10] и L1/GIoU Loss для DAB-DETR. Тестирование сетей проводилось с порогом уверенности (Confidence) 0,25 и порогом IoU для Non-Maximum Suppression (NMS) 0,6, что является более строгим критерием и позволяет снизить количество ложноположительных срабатываний.

Для определения качества работы обученных алгоритмов детекции использовались стандартные метрики mAP@50, mAP@50-95, F1, Precision (P), Recall (R) [11-12]. Метрика mAP@50-95 рассчитывается как среднее из 10-ти значений mAP при 10 различных порогах IoU в интервале от 0,50 до 0,95 с шагом 0,05. Функции потерь на этапе обучения и валидации в зависимости от номера эпохи, а так же графики зависимости mAP@50 и mAP@50-95 на этапе валидации от номера эпохи приведены на рисунке 2.

В ходе эксперимента было зафиксировано стабильное уменьшение функции потерь на всех архитектурах, что свидетельствует о корректности подобранных гиперпараметров. Наиболее быстрая сходимость наблюдалась у детекторов YOLOv11, тогда как архитектуры на трансформерах требовали больше эпох для выхода на «плато». Выбор финальных весов моделей для тестирования осуществлялся на основе эпохи с наилучшим значением метрики mAP@50-95 на валидационной выборке.



а) б)



в) г)

Рисунок 2 – Кривые обучения и валидации модели YOLOv11m: а) функции потерь на этапе тренировки; б) функции потерь на этапе валидации; в) динамика метрики mAP@50 на валидационной выборке; г) динамика метрики mAP@50-95 на валидационной выборке

Результаты тестирования обученных нейросетевых алгоритмов детектирования на тестовой выборке представлены в таблице 2.

Таблица 2. Результаты тестирования нейросетевых алгоритмов детектирования

Архитектура	**P**	**R**	**F1**	**mAP@50**	**mAP@50-95**
RTMDet	0,948	0,991	0,969	0,949	0,61
TOOD	0,963	0,982	0,973	0,964	0,674
DAB-DETR	0,966	0,984	0,975	0,974	0,679
YOLOv11s	0,992	0,995	0,993	0,992	0,756
YOLOv11m	0,992	0,995	0,993	0.994	0,755

Согласно результатам тестирования из таблицы 2 наилучший результат по всем ключевым метрикам продемонстрировали модели семейства YOLOv11 (версии s и m). Они достигли высоких значений Precision (0,992) и Recall (0,995), что отражается в максимальном F1 (0,993) и mAP@50-95 (0,756), которая оценивает качество детекций при различных порогах IoU. Архитектуры DAB-DETR и TOOD показали схожие и весьма достойные результаты, немного уступая YOLOv11 по точности, но превосходя RTMDet. RTMDet, несмотря на высокий Recall, имеет более низкие значения Precision и mAP@50-95, что может указывать на большее количество ложных срабатываний в сложных условиях. Дополнительно следует отметить, что модели YOLOv11 обеспечивают сбалансированное соотношение точности и полноты детекции, что важно при анализе эндоскопических видеоданных, содержащих большое количество визуально схожих структур. Высокие значения mAP@50-95 свидетельствуют о стабильности локализации объекта при различных порогах IoU, что повышает надежность применения модели в системах автоматического контроля качества колоноскопических исследований.

Заключение

Проведенное сравнительное исследование современных нейросетевых архитектур для детекции купола слепой кишки на эндоскопических изображениях подтвердило высокую эффективность подхода на основе глубокого обучения для решения данной клинически значимой задачи. Наилучшие результаты продемонстрировала архитектура YOLOv11, что делает её предпочтительным кандидатом для интеграции в систему автоматизированного контроля качества колоноскопических исследований. Для дальнейшего повышения надежности системы и снижения ложноположительных срабатываний в условиях высокой визуальной схожести КСК с фоновыми структурами кишечника предлагаются следующие направления – расширение обучающей выборки за счёт большего числа исследований и сложных случаев, целенаправленное дообучение на примерах ошибок, а также внедрение алгоритмов трекинга для анализа временной согласованности детекций в видеопотоке. Реализация модуля на основе предложенного решения позволит объективно контролировать полноту проведения процедуры колоноскопии, способствуя повышению стандартов диагностики.

Литература

1. Куваев Р.О., Никонов Е.Л., Кашин С.В., Капранов В.А., Гвоздев А.А. Контроль качества эндоскопических исследований, перспективы автоматизированного анализа эндоскопических изображений // Кремлевская медицина. Клинический вестник. 2013. Т. 2. С. 51–56.
2. Хрящев В.В., Завьялов Д.В., Андержанова А.С. Классификация эндоскопических изображений устья червеобразного отростка на основе методов глубокого машинного обучения // Цифровая обработка сигналов. 2023. № 1. С. 35-38.
3. Завьялов Д. В., Котов Н. В., Хрящев В. В. Исследование алгоритма оптического потока для анализа колоноскопических видеоданных // Измерение. Мониторинг. Управление. Контроль. 2025. № 4. С. 154–164.
4. Khanam R., Hussain M. Yolov11: An overview of the key architectural enhancements // arXiv preprint arXiv:2410.17725. 2024.
5. Wang C. Y. et al. CSPNet: A new backbone that can enhance learning capability of CNN // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020. P. 390-391.
6. Lyu C. et al. Rtmdet: An empirical study of designing real-time object detectors // arXiv preprint arXiv:2212.07784. 2022.
7. Feng C. et al. Tood: Task-aligned one-stage object detection // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society, 2021. pp. 3490-3499.
8. Liu S. et al. Dab-detr: Dynamic anchor boxes are better queries for detr // arXiv preprint arXiv:2201.12329. 2022.
9. Loshchilov I., Hutter F. Decoupled weight decay regularization // arXiv preprint arXiv:1711.05101. 2017.
10. Lin T. Y. et al. Focal loss for dense object detection // Proceedings of the IEEE international conference on computer vision. 2017. P. 2980-2988.
11. Jegham N. et al. Evaluating the evolution of yolo (you only look once) models: A comprehensive benchmark study of yolo11 and its predecessors // arXiv e-prints. 2024. S. arXiv: 2411.00201.
12. Хрящев В.В., Котов Н.В., Тихомиров А.А., Ненахов И.С. Комбинированный алгоритм анализа изображений для контроля качества колоноскопического исследования // Биомедицинская радиоэлектроника. 2025. Т. 28. № 1. С. 64-74.

COMPARATIVE ANALYSIS OF NEURAL NETWORK ARCHITECTURES FOR REGION OF INTEREST DETECTION IN COLONOSCOPY VIDEO DATA

Kotov N.V.

P.G. Demidov Yaroslavl State University

The problem of quality control of colonoscopic examinations based on neural network detection of the orifice dome area is considered. To solve this problem, modern detectors based on convolutional neural networks and transformers are being investigated – YOLOv11, RTMDet, TOOD and DAB-DETR. Computer modeling was performed on our own dataset containing 6927 pairs of images and markup files, as well as background images. According to the test results, the YOLOv11 architecture demonstrated the best results in terms of mAP@50, mAP@50-95, and F1-measure metrics, showing results of 0.994, 0.755, and 0.993, respectively. The study will serve as the basis for the development of a module for automatic monitoring of the completeness of colonoscopic examinations in real clinical conditions.

ПРЕДВАРИТЕЛЬНАЯ ФИЛЬТРАЦИЯ СПЕКЛ-ШУМА В ЗАДАЧЕ АНАЛИЗА SAR-ИЗОБРАЖЕНИЙ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМНОЙ ПОВЕРХНОСТИ

асп. Сенников А.В., асп. Ларионов Р.В., студ. Сидоров Н. А.

Ярославский государственный университет им. П.Г. Демидова

Широкое использование изображений синтезированной апертуры (SAR, РСА) в задачах дистанционного зондирования, таких как оценка биомассы лесов и прогнозирование стихийных бедствий, затруднено из-за присущего им спекл-шума. Этот мультипликативный, зависящий от сигнала шум ухудшает качество изображений и осложняет их анализ. В статье представлено исследование классических и современных алгоритмов на основе нейронных сетей для удаления спекл-шума с РСА-изображений. В работе сравнивается эффективность фильтра BM3D, улучшенного фильтра Ли и двух современных подходов на основе глубокого обучения: алгоритма SAR Transformer и алгоритма SAR2SAR. Оценка проводилась на тестовых изображениях, искаженных спекл-шумом. Производительность каждого алгоритма оценивалась с использованием комплексного набора метрик оценки качества, включающего MSE, PSNR, SSIM, ENL, SMPI и EPI, что позволило провести количественный анализ способностей подавления шума и сохранения деталей.

Введение

Данные дистанционного зондирования земной поверхности (ДЗЗ) нашли широкий спектр применений. Задачи в этой области включают, например, оценку биомассы лесов, управление сельскохозяйственными и природными ресурсами, геологию и прогнозирование стихийных бедствий [1]. Спутниковые данные, полученные в оптическом диапазоне часто не являются наиболее подходящим источником информации для таких задач, так как сильно зависят от погодных условий. Поэтому для задач подобного рода используются спутниковые снимки в радиодиапазоне, получаемые активными радиолокационными устройствами, такими как радары с синтезированной апертурой (SAR, РСА) [2].

Однако изображения, полученные с использованием РСА-систем, содержат высокий уровень шума из-за природы радиолокационного обратного рассеяния. Основным источником шума на РСА-изображениях является спекл-шум, который возникает в результате взаимодействия когерентного луча радара с исследуемой неровной поверхностью [3].

Моделирование спекл-шума

Спекл-шум проявляется как зернистая или пятнистая структура на изображениях, что затрудняет их интерпретацию и анализ. Природа спекл-шума заключается в том, что радиолокационные системы излучают когерентные микроволновые сигналы, которые отражаются от поверхности Земли и возвращаются к приемнику [2]. Затем эти отраженные волны складываются когерентно, что приводит к образованию интерференционной картины. В результате этой интерференции происходит случайное усиление или ослабление сигнала, создающее зернистую структуру.

Для моделирования используется мультипликативная модель шума, при которой исходное изображение умножается на модель случайного распределения с единичным математическим ожиданием [2, 4]. Таким распределением может быть распределение Рэлея, Гамма или нормальное распределение. В данной работе для моделирования используется нормальное распределение. Математически модель спекл-шума можно выразить следующим образом:

$$noised_signal(x) = signal(x) * S,$$

где $noised_signal$ – зашумленный сигнал, $signal$– исходный сигнал, S– нормальное случайное распределение.

Алгоритмы фильтрация спекл-шума и оценки качества изображений

Возможность эффективного удаления спекл-шума с изображений имеет важное значение в задачах ДЗЗ. Исследование изучает и сравнивает эффективность четырех современных подходов к удалению спекл-шума, представляющих разные классы алгоритмов: от классических нелокальных и адаптивных фильтров до современных архитектур нейронных сетей.

BM3D (Block-Matching and 3D Filtering) – нелокальный фильтр, основанный на совместной фильтрации. Его алгоритм включает два основных шага: сначала выполняется группировка похожих патчей изображения в 3D-блоки и их фильтрация с использованием пороговой обработки, затем выполняется второй шаг фильтрации Винера с использованием базовой оценки, полученной на первом этапе [5].

Улучшенный фильтр Ли – это адаптивный фильтр, который подавляет шум путем взвешенного усреднения в локальном окне на основе статистики яркости: в однородных областях применяется сильное сглаживание, а вблизи границ и деталей фильтрации ослабевает для их сохранения [6].

SAR Transformer – это нейросетевой фильтр с гибридной архитектурой, использующий механизм внимания для анализа глобального контекста изображения, что позволяет эффективно отделять полезный сигнал от спекл-шума [7].

SAR2SAR является полуавтоматическим нейросетевым фильтром, который обучается на парах зашумленных изображений, что устраняет необходимость в чистых эталонах, и использует адаптированную функцию потерь для эффективной работы с реальными PCA-данными [8].

Оценка производительности исследуемых фильтров проводится с использованием стандартных метрик качества, применяемых к тестовым изображениям, искаженным спекл-шумом. Оценка включает как универсальные метрики качества изображений, так и специальные индексы для анализа PCA-изображений [9].

Основными количественными метриками для оценки подавления шума являются MSE (Mean Square Error) и PSNR (Peak Signal-to-Noise Ratio). Метрика MSE измеряет среднюю квадратичную разницу между исходным изображением и отфильтрованным изображением [10]:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}|I(i,j) - K(i,j)|^2.$$

Метрика PSNR выражает отношение между максимально возможной мощностью сигнала и мощностью искажаемого шума [10]:

$$PSNR = 20\lg\left(\frac{MAX_I}{\sqrt{MSE}}\right).$$

Для оценки сохранения структурной информации используется метрика SSIM (Structural Similarity Index). Он сравнивает яркость, контраст и структуру между двумя изображениями [10].

Для специфического анализа PCA-изображений важной метрикой является ENL (Equivalent Number of Looks). Метрика оценивает уровень сглаживания, достигнутый в однородных областях, путём вычисления отношения квадрата среднего значения к дисперсии. Более высокое значение ENL указывает на лучшее подавление спекл-шума [10]:

$$ENL = \left(\frac{mean}{std}\right)^2.$$

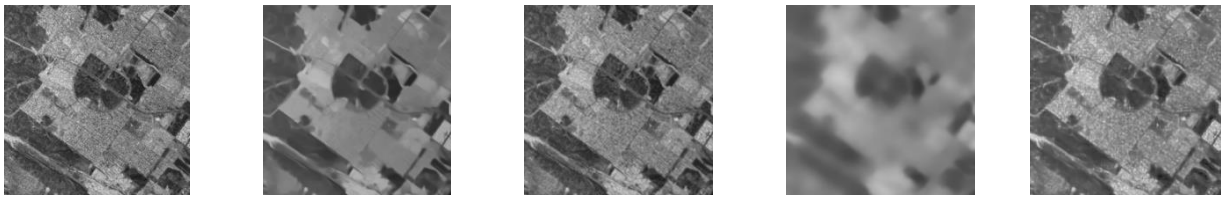
Метрика SMPI (Speckle Suppression and Mean Preservation Index) предоставляет сбалансированную меру, учитывающую как снижение шума, так и сохранение среднего радиометрического значения. Более низкое значение SMPI указывает на лучшую общую производительность [11]:

$$SMPI = \left[1 + \left|mean(I_0) - mean(I_f)\right|\right]\left(\frac{std(I_f)}{std(I_0)}\right).$$

Метрика EPI (Edge Preservation Index) количественно оценивает способность фильтра сохранять мелкие детали и границы объектов. EPI применяет детектор границ Кэнни как к исходному, так и к отфильтрованному изображению, преобразует результаты в бинарные маски [12].

Результаты моделирования

Оценка четырех ранее описанных алгоритмов удаления спекл-шума: BM3D, улучшенного фильтра Ли, SAR Transformer и SAR2SAR проводилась на разнообразном наборе тестовых изображений, искажённых спекл-шумом для имитации реальных условий РСА-съемки. Полученные результаты приведены на рисунке 1 и таблицах 1 и 2.



а) б) в) г) д)

Рисунок 1 – Результаты тестирования на изображении «San Diego»: а) изображение с шумом при var = 0,1 и обработка фильтрами: б) BM3D, в) lee Enh., г) SAR Tr., д) SAR2SAR

Таблица 1. Значения полученных метрик оценки качества для изображения «San Diego»

var	метрика	Шум	BM3D	Lee Enh.	SAR Tr.	SAR2SAR
0,1	PSNR	26,58	24,18	25,32	18,46	24,40
	SSIM	0,83	0,58	0,74	0,19	0,77
	ENL	17,91	17,91	28,31	431,15	32,27
	SMPI	1,61	1,61	1,52	3,26	8,08
	EPI	0,64	0,64	0,49	0,01	0,53
0,3	PSNR	17,33	23,21	21,63	18,08	16,55
	SSIM	0,41	0,59	0,54	0,18	0,49
	ENL	7,38	36,49	18,28	476,05	15,73
	SMPI	4,39	2,01	2,80	3,31	23,41
	EPI	0,44	0,39	0,40	0,01	0,43

Таблица 2. Значения полученных метрик оценки качества для изображения «Field»

var	метрика	Шум	BM3D	Lee Enh.	SAR Tr.	SAR2SAR
0,1	PSNR	31,29	19,20	19,63	15,76	14,21
	SSIM	0,98	0,40	0,64	0,21	0,54
	ENL	2,22	37,92	5,96	56,58	8,05
	SMPI	1,30	0,67	0,99	2,03	14,37
	EPI	0,86	0,33	0,53	0,19	0,57
0,3	PSNR	21,85	19,43	19,26	16,23	13,11
	SSIM	0,87	0,48	0,63	0,29	0,49
	ENL	1,79	18,23	4,71	28,79	5,59
	SMPI	2,05	1,12	1,37	2,92	20,85
	EPI	0,66	0,40	0,51	0,37	0,52

Проведенный сравнительный анализ четырех алгоритмов подавления спекл-шума (BM3D, улучшенный фильтр Ли, SAR Transformer, SAR2SAR) выявил их сильные и слабые стороны. Классические методы (BM3D, фильтр Ли) обеспечивают стабильное шумоподавление, но зачастую в ущерб деталям и границам объектов. Нейросетевые подходы показали разную эффективность. SAR Transformer демонстрирует выдающееся подавление шума, однако чрезмерно сглаживает изображение, теряя важные детали. Напротив, SAR2SAR, обучающийся на парах зашумленных данных, предлагает лучший баланс между удалением шума и сохранением структурной информации.

Заключение

Таким образом, выбор алгоритма зависит от конкретной задачи. Для максимального подавления шума подходят BM3D или SAR Transformer, для быстрой предобработки — фильтр

Ли, а для практических задач ДЗЗ, требующих сохранения деталей, наиболее перспективны адаптивные нейросетевые методы, подобные алгоритму SAR2SAR.

Литература

1. Рис У.Г. Основы дистанционного зондирования. – М.: Техносфера, 2006. – 336 с.
2. Гук А.П., Евстратова Л.Г. Дистанционное зондирование и мониторинг территорий. Ч. 1. Дистанционное зондирование. Теоретические основы и технические средства: Учебное пособие. – М.: КУРС, 2025. – 224 с.
3. Maity A. A comparative study on approaches to speckle noise reduction in images. / 2015 International Conference on Computational Intelligence and Networks. – IEEE, 2015. – pp. 148–155.
4. Сенников А.В., Ларионов Р.В., Хрящев В.В. Сравнительный анализ использования классических и нейросетевых алгоритмов для фильтрации спекл-шума из изображений // Цифровая обработка сигналов и ее применение (DSPA-2025): докл. 27-й междунар. конф. Москва, 2025. С. 192–195.
5. Yahya A. A. и др. BM3D image denoising algorithm based on an adaptive filtering // Multimedia Tools and Applications. – 2020. – Т. 79. № 27. – P. 20391–20427.
6. Старовойтов В. В. Методика выбора фильтра для сглаживания спекл-шума радарных изображений с синтезированной апертурой / Обработка сигналов, изображений и речи. – 2015. – С. 5–9.
7. Perera M. V., Bandara W. G. C., Valanarasu J. M. J., Patel V. M. Transformer-Based SAR Image Despeckling // International Geoscience and Remote Sensing Symposium (IGARSS), 2022. – P. 751-754.
8. Dalsasso E., Denis L., Tupin F. SAR2SAR: a semi-supervised despeckling algorithm for SAR images // arXiv:2006.15037v3 [eess.IV] 13 Apr 2021.
9. Larionov R., Sennikov A., Khryashchev V., Priorov A. Quality Image Assessment after Despeckling in Remote Sensing // IEEE XVII International Conference on Actual Problems of Electronic Instrument Engineering (APEIE-2025), Penza, 2025.
10. Приоров А. Л., Хрящев В.В., Топников А.И. Обработка и передача мультимедийной информации: учебное пособие – Ярославль, ЯрГУ, 2022 – 248 с.
11. Sankar S. A comparative analysis of speckle filtering in SAR images // International Conference on Signal Processing and Communications (SPCOM). – 2016. – pp. 1–5.
12. Lee J.-S., Grunes M.R., Schuler D.L. Polarimetric SAR speckle filtering and its implication for classification – IEEE, 1999. – pp. 2363–2373.

PRE-FILTERING SPECKLE NOISE IN EARTH SURFACE SAR IMAGE ANALYSIS

A.V. Sennikov, R.V. Larionov, N.A. Sidorov

P.G. Demidov Yaroslavl State University

The widespread use of synthetic aperture radar (SAR) images in remote sensing applications, such as forest biomass assessment and disaster forecasting, is hampered by their inherent speckle noise. This multiplicative, signal-dependent noise degrades image quality and complicates their analysis. This paper presents a study of classical and modern neural network-based algorithms for speckle noise removal from SAR images. This paper compares the performance of the BM3D filter, an improved Lee filter, and two modern deep learning-based approaches: the SAR Transformer algorithm and the SAR2SAR algorithm. The evaluation was conducted on test images corrupted by speckle noise. The performance of each algorithm was assessed using a comprehensive set of quality assessment metrics, including MSE, PSNR, SSIM, ENL, SMPI, and EPI, allowing for a quantitative analysis of their noise reduction and detail preservation capabilities.

СРАВНЕНИЕ ЭФФЕКТИВНОСТИ НЕКОТОРЫХ НЕЙРОСЕТЕВЫХ МУЛЬТИМОДАЛЬНЫХ МОДЕЛЕЙ ДЛЯ ВИЗУАЛЬНОЙ ПРИВЯЗКИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ (НА ПРИМЕРЕ ДОМЕНА КУХОННОЙ УТВАРИ)

Калушев К.А.

ООО «Научно-производственное предприятие «Аквасинт»

Совершенствование мультимодальных моделей, происходящее в последние годы, открыло новые практические горизонты для решения задач в области компьютерного зрения. Одной из таких вновь появившихся задач выступает визуальная привязка (visual grounding) объектов, то есть локализация объектов определенного класса на изображении по текстовому описанию. В настоящей работе с учётом принципа «из коробки» проводится исследование широко известных в научной литературе мультимодальных моделей OWL-ViT, Grounding DINO, YOLO-World, Florence-2 с точки зрения эффективности визуальной привязки в домене кухонной утвари. Научная новизна работы заключается в оценке возможности практического использования моделей не на публичном наборе данных, а на пользовательском наборе данных, имитирующем реальные условия возможного применения в бытовой сервисной робототехнике. Для оценки эффективности моделей использован показатель accuracy@0.5, описана методика проведения эксперимента и продемонстрировано, что на настоящий момент в домене кухонной утвари передовые нейросетевые модели обеспечивают аккуратность визуальной привязки объектов в диапазоне от 54 до 72%, а предельная эффективность возможного ансамбля моделей не превысит 88%.

Введение

Интеграция технологий обработки естественного языка и компьютерного зрения сделала шаг от способных к генерации только текста больших языковых моделей [1] к развитию мультимодальных моделей, интерпретирующих и связывающих информацию, полученную с помощью различных сенсорных средств, прежде всего, камер, с текстовыми описаниями окружающего мира.

Основанная на использовании мультимодальных моделей визуальная привязка (англ. – visual grounding) представляет собой задачу увязки лингвистических отсылок (текстов) с соответствующими визуальными объектами на изображении и является важнейшим компонентом в разработке интеллектуальных систем, способных понимать окружающую среду и взаимодействовать с ней. Актуальным практическим применением визуальной привязки выступает бытовая сервисная робототехника, на ее основе возможно создание робота, который «понимает», «видит» окружающее пространство и может выделить из него конкретные объекты.

Выбор для исследования визуальной привязки именно в домене кухонной утвари обусловлен (а) практической направленностью на решение таких задач как использование умных помощников на кухне, автоматизированное следование рецептам и управление запасами, (б) желанием отойти от традиционных наборов данных и (в) пониманием того, что кухонная обстановка характеризуется разнообразием предметов различной формы и размеров.

Цель настоящей работы – провести сравнительный анализ различных мультимодальных моделей для локализации на фотографических изображениях объектов, относящихся к домену кухонной утвари и соответствующих принадлежностей, на основании их текстовых описаний. Исследование проводится на авторском наборе тестовых изображений. Приводится сравнительная эффективность мультимодальных моделей на принципе «из коробки» (без дообучения под домен кухонной утвари), описывается методика и используемые метрики. Новизна исследования заключается как в домене используемых изображений, так и в составе моделей, в отношении которых проводится исследование.

Обзор литературы

В самом широком смысле под общей задачей компьютерного зрения по определению объектов можно понимать установление классов объектов на изображениях и их локализацию в пределах таких изображений. Вместе с тем, постановка данной задачи на современном этапе имеет множество практических направлений и менялась по мере развития технологий компьютерного зрения.

Ранние методы компьютерного зрения, такие как метод Виолы-Джонса (англ. – Viola-Jones detector) [2] и метод гистограмм ориентированных векторов (англ. – Histogram of Oriented Gradients (HOG)) [3], использовали прямо запрограммированные способы извлечения характерных признаков, позволяющих определять объекты и их положение на изображениях. Эти методы, что главное в контексте настоящей работы, определяли объекты только заранее заданных классов (и, тем более, не могли произвольно определять отдельные объекты по характерным только для них чертам, например, положению на изображении относительно других объектов).

Следующий шаг в эволюции общей задачи определения объектов на изображениях был связан как с появлением крупных наборов данных, таких как ImageNet [4], PASCAL VOC [5], MS COCO [6] , and Open Images [7] , так и с развитием сверточных нейронных сетей, приведшему к созданию ставшими классическими моделей Faster R-CNN [8] и SSD [9]. При этом обучение моделей осуществлялось на известных наборах данных с некоторым спектром заранее определенных категорий. Относительно незначительное количество таких категорий и представляло собой самое узкое место в использовании классических детекторов объектов на практике, например, 20 классов в Pascal VOC, 80 классов in COCO, и даже один из крупнейших наборов данных LVIS [10] включает только 1,203 категории.

С развитием визуально-языковых моделей (англ. – Vision-Language Models) решение широкой задачи определения объектов на изображениях вступило в новую эру – обнаружения объектов, относящихся к классам, которые не были напрямую заданы при обучении моделей, то есть без предварительного обучения. Конкретно визуальная привязка референтных выражений (англ. – Visual grounding of referring expressions или Referring Expression Comprehension, или Phrase Grounding) в классическом понимании, как например в [11] , заключается в том, чтобы для данного изображения и теста описательного характера, относящегося к объекту, как правило, включающему общий контекст сцены, определить объект, на который ссылается текст, и локализовать его.

Для сравнения в настоящем исследовании были выбраны следующие мультимодальные модели, которые достаточно часто упоминаются в работах на тему визуальной привязки, однако имеют различные архитектуры:

1. OWL-ViT, представленная в работе [12], является простым и масштабируемым детектором объектов с открытым словарем, использующим в своей основе визуальные трансформеры (англ. – Vision Transformer),
2. Grounding DINO, представленная в работе [13] использует свин-трансформеры (англ. – Swin Transformer) в качестве визуального энкодера,
3. YOLO-World, представленная в работе [14], использующая стандартную архитектуру YOLO (на базе YOLOv8) и репараметризацию для соединения сверточной нейронной сети с текстовыми эмбеддингами,
4. Florence-2, представленная в работе [15], в отличие от предыдущих трех моделей является фундаментальной (англ. Foundational) визуально-языковой моделью, нацеленной на решение не только задачи визуальной привязки объектов, но и целого ряда иных задач в области компьютерного зрения (включая, целостное описание изображений, семантическую сегментацию и распознавание текстов на изображениях).

Постановка задачи

Пусть имеется фотографическое изображение $I \in R^{H \times W \times 3}$, где H – высота изображения, а W – ширина изображения. На I присутствует эталонный набор из J объектов из домена кухонной утвари $O_{эт} = \{(k_j, b_j)\}$, где k_j – класс j-ого объекта (из общего количества классов K), а b_j – соответствующие эталонные ограничивающие рамки (англ. – bounding boxes), размеченные исследователем.

Пусть существует отсылочное выражение на естественном языке L, состоящее из T слов w_t, такое что $L = (w_1, w_2, \dots, w_t)$. При этом каждое w_t принадлежит некоторому словарю V, который должен включать k_j.

Пусть существует модель визуальной привязки M_g, которая может применяться к I в контексте L.

Тогда под задачей визуальной привязки подразумевается генерация ограничивающих рамок $B' = \{b'_i\}$, таких что $B' = M_g(I, L)$. При этом в классическом случае $i = 1$, то есть предполагается наличие только одного объекта, относящегося к заданному L. Оценка эффективности модели основана на том, насколько $b'_i \cong b_j \mid k_j \in K$.

Набор данных и методика проводимого эксперимента

Практическая цель исследования – определение возможности использования «из коробки» существующих популярных нейросетевых моделей для визуальной привязки объектов из домена кухонной утвари. Эксперимент был осуществлен в следующем порядке:

1. Выбраны 4 модели (OWL-ViT, Grounding DINO, YOLO-World, Florence-2).
2. Подготовлен пользовательский набор данных (изображения в формате JPG), который разделен на два суб-домена (две принципиально разные сцены – различные фоны, освещения и наборы объектов) с 73 различными классами (например, «тарелка с огурцами», «чайник», «электрическая розетка»), всего 214 различных объектов, относящихся к указанным классам.
3. Для каждого изображения исследователем размечаются эталонные ограничивающие рамки, соответствующие объектам присутствующих на изображении классов (общее количество уникальных классов на i-ой фотографии равно k_i) и присваиваются наименования классов (в связи с особенностями мультимодальных моделей наименования присваивались на английском языке). Разметка осуществляется с помощью программного продукта labelImg в формате PascalVOC. Пример разметки показан на рисунке 1.



(а) (б)

Рисунок 1 – Пример разметки изображения: (а) изображение до выделения границ объектов соответствующих классов, (б) изображение после выделения границ соответствующих классов и подписи классов

4. Все классы на каком-либо отдельном изображении делятся на две категории:

- классы, которые представлены на данном конкретном изображении только одном объектом (одним экземпляром), обозначаемые k’,
- классы, которые представлены на данном конкретном изображении в количестве двух и более объектов (например, на изображении несколько «чашек»).

Эксперименты по визуальной привязке осуществлялись только с классами, представленным на данном изображении в одном экземпляре (всего 120 объектов, относящихся к таким классам).

5. Для каждой модели для каждого i-ого изображения формируется k_i’ запросов «*Locate <object>*», где переменная *<object>* представляет собой наименование каждого уникального класса для данного изображения.
6. Для каждого класса успешной привязкой признавалась ситуация, когда на фотографическом изображении моделью определялся предполагаемый объект данного класса и

показатель IoU (англ. – Intersection over Union) для пар ограничивающих рамок привязанного объекта и эталонного объекта (ранее размеченного) превышал 0.5. Показатель IoU определялся по следующей формуле:

$$IoU^{\text{дет x эт}} = \frac{|Im_{\text{дет}} \cap Im_{\text{эт}}|}{|Im_{\text{дет}} \cup Im_{\text{эт}}|},$$

где $Im_{дет}$ – площадь прямоугольника, ограниченного рамками, определенными моделью, $Im_{эт}$ – площадь прямоугольника, ограниченного эталонными рамками, определенными исследователем заранее.

На рисунке 2 представлены примеры успешной и неуспешной привязки объектов.



(а) (б)

Рисунок 2 – Пример (а) неуспешной привязки объекта класса «Sugar bowl» (сахарница) и (б) успешной привязки объекта класса «Cutting board» (доска для нарезки) на различных изображениях

7. После определения успешных/неуспешных привязок объектов для каждого изображения эффективность моделей оценивалась с помощью метрики accuracy@0.5, которая использовалась, например, с сопоставительным исследованием [16], по формуле:

$$accuracy@0.5 = \frac{M \mid IoU \geq 0.5}{N}$$

где N – общее эталонное количество объектов всех k’ классов на I, определенное исследователем, а M_i – количество объектов из k’, определенное соответствующей моделью, при условии, что IoU превышает пороговое значение в 0.5.

Результаты экспериментов

Результаты экспериментов по привязки объектов различных классов на изображениях в домене кухонной утвари приведены в таблице 1.

Таблица 1. Результаты экспериментов по привязке объектов для различных нейросетевых моделей

Количество объектов в домене/суб-домене	68	52	120
	Суб-домен №1 Accuracy@0.5	*Суб-домен №2 Accuracy@0.5*	*Домен целиком Accuracy@0.5*
Florence-2	54,4%	59,6%	56,7%
OWV-ViT	64,7%	80,8%	71,7%
GroundingDINO	64,7%	59,6%	62,5%
YOLO-World	60,3%	46,2%	54,2%
Любая модель	85,3%	92,3%	88,3%

Эффективность мультимодальных моделей для визуальной привязки объектов на домене кухонной утвари по принципу «из коробки» при нахождении одного целевого объекта на изображении находится в диапазоне от 54 до 72% (в среднем, около 60%). При этом, как продемонстрировано, эффективность зависит от суб-домена в связи с тем, что в суб-домене №1 показатель accuracy@0.5 практически одинаков для всех моделей, в то время как в суб-домене №2 разброс его значений существенен.

В таблицу 1 также включена строка «Любая модель», которая соответствует ситуации, когда для данного класса хотя бы одна модель осуществила визуальную привязку точно. В силу очевидных причин данный показатель выше, чем у какой-либо отдельно взятой модели, составляя для домена в целом 88%. Вместе с тем, факт того, что и он не достигает близкого к стопроцентному значения, демонстрирует, что даже ансамбль из моделей на сегодняшний день не может полностью решить задачу визуальной привязки на реальных пользовательских наборах данных, как минимум в домене кухонной утвари.

Выводы

Визуальная привязка является одной из новейших задач компьютерного зрения, результаты решения которой потенциально могут быть непосредственно использованы на практике, прежде всего, в робототехнике для организации диалога машины и оператора. Настоящая работа была призвана продемонстрировать эффективность решения данной задачи различными широко известными мультимодальными моделями не на публичном наборе данных, а на специально созданном для демонстрации возможности решения практических задач в области сервисной бытовой робототехники (кухонной утвари).

Проведенное исследование продемонстрировало, что диапазон аккуратности мультимодальных моделей по принципу «из коробки» (то есть без донастройки и использования ансамблей моделей) для домена кухонной утвари составляет от 54 до 72%, при этом значительно зависит от суб-домена, но даже для теоретического ансамбля моделей не превысит 88%.

Представляется, что несмотря на совершенный в последнее время мультимодальными моделями прорыв, уровень их точности «из коробки» все еще не достаточен для практического применения в робототехнике за рамками лабораторных исследований, открывая целый горизонт для совершенствования как принципиальных архитектур моделей.

Литература

1. Лещев, С. В. Генеративная мысль: преподавание в эпоху больших языковых моделей / С. В. Лещев // Современные проблемы гуманитарных и общественных наук. – 2025. – № 4(56). – С. 41-45. – EDN ZPIRRL.

2. Viola, P. Rapid object detection using a boosted cascade of simple features, in Proc / P. Viola, M. Jones // IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. CVPR, vol. 1, Apr. 2001, pp. 1511–1518.

3. Surasak, T. Histogram of oriented gradients for human detection in the video / T. Surasak, I. Takahiro, and C-H. Cheng // in Proc. 5th Int. Conf. Bus. Ind. Res. (ICBIR), 2018, pp. 172–176.

4. Deng, J. ImageNet: A large-scale hierarchical image database / J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei // in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 248–255.

5. Everingham, M. The PASCAL visual object classes challenge 2012 (VOC2012) development kit / M. Everingham, J. Winn // Pattern Anal. Stat. Model. Comput. Learn., vol. 2007, Jan. 2012, pp. 1–45.

6. Lin, T. Microsoft COCO: Common objects in context / T. Lin // in Proc. Eur. Conf. Comput. Vis., 2014, pp. 740–755.

7. Kuznetsova, A. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale / A. Kuznetsova, H. Rom, N. Alldrin, J. Uijlings, I. Krasin, J. Pont-Tuset, S. Kamali, S. Popov, M. Malloci, A. Kolesnikov, T. Duerig, and V. Ferrari // Int. J. Comput. Vis., vol. 128, no. 7, Jul. 2020, pp. 1956–1981.

8. Shaoqing, R. Faster R-CNN: towards real-time object detection with region proposal networks / R. Shaoqing, H. Kaiming, G. Ross, S. Jian // In Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'15), Vol. 1. MIT Press, Cambridge, MA, USA, 2015, 91–99.

9. Liu, W. SSD: Single Shot MultiBox Detector / W. Liu, B. Leibe, J. Matas, N. Sebe, M. Welling, // Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, vol 9905, 2016, Springer, pp 21–37.

10. Gupta, A. LVIS: A dataset for large vocabulary instance segmentation / A. Gupta, P. Dollar, and R. Girshick // in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019, pp. 5351–5359.

11. Deng, J. Transvg: End-to-end visual grounding with transformers / J. Deng, Z. Yang, T. Chen, W. Zhou, and H. Li // in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 1769–1779.

12. Minderer, M. Simple Open-Vocabulary Object Detection / M. Minderer, S. Avidan, G. Brostow, M. Cissé, G.M. Farinella, T. Hassner // Computer Vision – ECCV 2022. ECCV 2022. Lecture Notes in Computer Science, vol 13670. Springer, 2022, pp 728–755.

13. Liu, S. Grounding DINO: Marrying DINO with Grounded Pre-training for Open-Set Object Detection / S. Liu, A. Leonardis, E.Ricci, S. Roth, O. Russakovsky, T. Sattler, G. Varol // Computer Vision – ECCV 2024. ECCV 2024. Lecture Notes in Computer Science, vol 15105. Springer, 2025, pp 38–55.

14. Tianheng, C. YOLO-World: Real-Time Open-Vocabulary Object Detection / C. Tianheng, C. Tianheng, S. Lin, G. Yixiao, L. Wenyu, W. Xinggang, S. Ying // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June, 2024, pp 16901–16911.

15. Xiao, B. Florence-2: Advancing a Unified Representation for a Variety of Vision Tasks / B. Xiao, H. Wu, W. Xu, X. Dai, H. Hu, Y. Lu, M. Zeng, C. Liu, L. Yuan // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June, 2024, pp 4818-4829.

16. Wang, J. A Simple and Better Baseline for Visual Grounding / J. Wang, W. Zhang, D. Huang, H. Wang and Y. Zheng // 2025 IEEE International Conference on Multimedia and Expo (ICME), Nantes, France, 2025, pp. 1-6.

COMPARISON OF THE EFFECTIVENESS OF SOME NEURAL NETWORK MULTIMODAL MODELS FOR VISUAL ANCHORING OBJECTS IN IMAGES (USING THE KITCHEN UTENSILS DOMAIN AS AN EXAMPLE)

Kalushev K.A.

OOO Industrial-Research Enterprise “Akvasint”

The improvement of multimodal models that has taken place in recent years has opened up new practical horizons for solving problems in the field of computer vision. One of these newly emerged tasks is the visual grounding of objects, that is, the localization of objects of a certain class in an image based on a text description. In this paper, taking into account the out-of-the-box principle, we study the multimodal OWL-ViT, Grounding DINO, YOLO-World, and Florence-2 models widely known in the scientific literature in terms of the effectiveness of visual binding in the domain of kitchen utensils. The scientific novelty of the work is to evaluate the possibility of practical use of models not on a public dataset, but on a user dataset that simulates the real conditions of possible use in household service robotics. The indicator was used to evaluate the effectiveness of the models accuracy@0.5 The method of conducting the experiment is described and it is demonstrated that at the moment, in the domain of kitchen utensils, advanced neural network models ensure the accuracy of visual binding of objects in the range from 54 to 72%, and the maximum efficiency of a possible ensemble of models will not exceed 88%.

ИСПОЛЬЗОВАНИЕ МУЛЬТИМОДАЛЬНЫХ МОДЕЛЕЙ В ОПРЕДЕЛЕНИИ ЖАНРОВ ФИЛЬМОВ

асп. Шувариков М.Д.

Финансовый университет при Правительстве Российской Федерации

Использование различных модальностей данных позволяет успешно повышать качество моделей. Одной из сфер, где растет интерес бизнеса к использованию данного подхода, является предсказание жанров фильмов на стриминговых сервисах, число пользователей которых особенно быстро увеличивается в последние годы. Цель данной работы – исследовать уже существующие подходы к прогнозированию жанров кино и предложить свою наиболее эффективную мультимодальную модель, обученную на данных с российского сервиса Кинопоиск.

Введение

В 2024 году число пользователей стриминговых платформ, таких как Netflix, Hulu, Prime Video, достигло 1,42 миллиарда зрителей [1], и ожидается, что в ближайшие годы этот показатель будет расти и достигнет 1,65 миллиарда в 2027 году. Данная тенденция подсвечивает необходимость расширения существующих технологий классификации жанров фильмов, поскольку это может помочь в улучшении действующих рекомендательных систем.

Фильм, как продукт, обычно идет «в комплекте» с различными вспомогательными материалами, такими как описание сюжета (синопсис), постер, трейлер, эта особенность предоставляет нам различные типы модальностей (типы данных), соответственно, мы можем попытаться использовать некоторые из них для создания высококачественной модели. Попытки работать с различными методами предпринимаются уже более 30 лет. Однако с развитием технологий и методик, таких как LLM (большие языковые модели), в 2020-х годах исследованиям в этой области снова стало уделяться особое внимание. С учетом того, что релевантных исследований данной темы в российских источниках и по российским данным нет, данная работа будет крайне полезна не только российскому сегменту стриминговых платформ, но и будущим отечественным исследователям.

Анализ предыдущих исследований

В работах 1990-х годов исследователи обычно использовали видео- и аудиоанализ. Например, Nam и Tewfik [2] пытались применить данные методы для разделения видеоряда на отдельные кадры. Другая группа ученых [3] использовала анализ аудиосигналов, чтобы выделить признаки (фичи) и классифицировать различные программы на телевидении – рекламные ролики, баскетбол, новости и т.д. В качестве непосредственно классификационной модели использовали k-ближайших соседей (kNN).

Одна из наиболее ранних работ, в которой Brezeale and Cook [4] использовали субтитры и фильмы (как набор кадров), была выпущена в 2006 году. Для извлечения признаков из первой модели был использован мешок слов (BOW), а для второй - мешок визуальных признаков (BOVF), для модели классификатора был выбран метод опорных векторов (SVM). Наилучший показатель качества (в данной работе – accuracy) составил 89,7%.

Jain and Jadon [5] сфокусировались на извлечении аудио- и видеофрагментов, используя такие фичи, как высота звука, частота, MFCC (Мел-частотные кепстральные коэффициенты) для первого типа данных и длина кадра, доминирование цвета, тональность освещения для второго. В качестве классификатора ученые выбрали многослойную нейронную сеть с прямой связью, которая продемонстрировала устойчивые результаты после объединения аудио и видео фичей в одну модель.

В более поздних работах 2010-х годов обычно использовались такие модальности (и их комбинации), как трейлеры, кадры, теги, синопсисы и сюжеты фильмов [6-7]. Средства извлечения фичей были представлены такими инструментами, как ранее упомянутые BOWs, BOVFs, TF-IDF, однако в конце 2010-х годов более современные инструменты, такие как CNN и LSTM, стали использоваться как в качестве дескриптора контента, так и в качестве модели [8-9]. Помимо нейронных сетей, в основном использовались алгоритмы SVM и деревья решений. Показатели также варьировались от автора к автору, но чаще использовался набор из следующих: Accuracy, F1 и PR-AUC.

Одно из наиболее релевантных исследований было проведено в 2020 году Mangolin et al. [10]. Авторы использовали такие модальности, как аудио, видеокадры, постеры, субтитры и

синопсис. Были применены такие средства извлечения фичей: MFCC (для аудио), сверточная 3D нейронная сеть (для видео), Local binary patterns (видео и постеры), N-граммы с TF-IDF (субтитры и синопсисы). В качестве моделей авторы выбрали kNN, LSTM, MLP, которые были объединены для расчета прогноза. В этом исследовании в качестве целевых метрик качества были выбраны PR-AUC и F1-score. Наилучшие результаты были получены при смешивании моделей для трейлера и синопсиса (0,628 F1-score), а также при смешивании моделей для трейлера, синопсиса и субтитров (0,673 PR-AUC).

Идеи совместной обработки изображений, содержащих текстовые вставки [11] также позволили получить выигрыш по классификации рекламных, поздравительных и «мемных» изображений.

Описанные работы наглядно демонстрируют необходимость использования нескольких модальностей, поскольку это может улучшить качество прогнозов.

Данные и методология исследования

Данные для этого исследования были собраны из российской онлайн-базы данных Кинопоиск [12], содержащей всю необходимую информацию о фильмах, такую как режиссеры, актерский состав, постеры, синопсисы и т.д. Мы создали датасет из 8335 фильмов в 14 основных жанрах, от драмы и комедии до фэнтези и детектива. Для каждого фильма мы взяли синопсис и постер в качестве двух источников фичей для будущей модели. Одной из отличительных особенностей таких онлайн-агрегаторов (как и самого фильма) является его многожанровость, то есть для каждой картины зрителю предоставляется несколько (на Кинопоиске обычно 3) жанров. Следовательно, мы будем работать и моделировать переменную с несколькими метками (multi-label variable).

Таблица 1. Распределение классов

Жанр	Число	Доля
Боевик	1406	16,9%
Вестерн	423	5,1%
Военный	504	6,0%
Детектив	717	8,6%
Драма	4705	56,4%
Комедия	3068	36,8%
Криминал	1643	19,7%
Мелодрама	2137	25,6%
Приключения	1050	12,6%
Семейный	480	5,8%
Триллер	1956	23,5%
Ужасы	710	8,5%
Фантастика	673	8,1%
Фэнтези	575	6,9%

Согласно данным, представленным в таблице 1, основными жанрами являются драма, комедия и мелодрама, в то время как наименьшая доля приходится на военные фильмы, вестерны и семейные. После сбора данные были случайным образом разделены на обучающую и валидационную выборки, с долями 80% и 20% соответственно.

Для решения нашей задачи мы решили провести fine-tuning (дообучение) предварительно обученных (pre-trained) текстовых и картиночных моделей, чтобы затем объединить их в одну модель, а после добавить и обучить «голову» классификатора на собранных выборках данных. Для pre-trained моделей мы выбрали следующие:

• vit-huge-patch14-224-in21k. Создана и обучена Google на наборе данных ImageNet-21k Vision Transformer. Модели семейства VIT высоко оцениваются в бенчмарке cifar-10, например, VIT-H/14 сейчас занимает первое место с точностью 99,5%. [13]

• rdnet_large. Плотно связанные сверточные сети (DenseNets) были обновлены командой Naver AI, поэтому эта модель входит в топ-10 классификаторов в бенчмарке cifar-10. [14]

Для классификации текстов мы выбрали такие модели:

• sbert_large_mt_nlu_ru. Эта модель Sentense BERT была создана командой AI-Forever и предназначена для решение задач на русском языке [15].

• rubert-base-cased. Это модель RuBERT, обученная на русском сегменте Википедии и новостях [16].

• multilingual-e5-base. Модель многоязычных текстовых эмбеддингов Microsoft может быть использована для русскоязычных задач [17].

Таким образом, ключевая идея создания нашего мультимодального классификатора заключается в извлечении текстовых и картиночных эмбеддингов с использованием ранее описанных моделей, объединении этих фичей, а затем отправке данных в голову классификатора, который будет обучен на собранных данных. Архитектура модели представлена на рисунке 1.

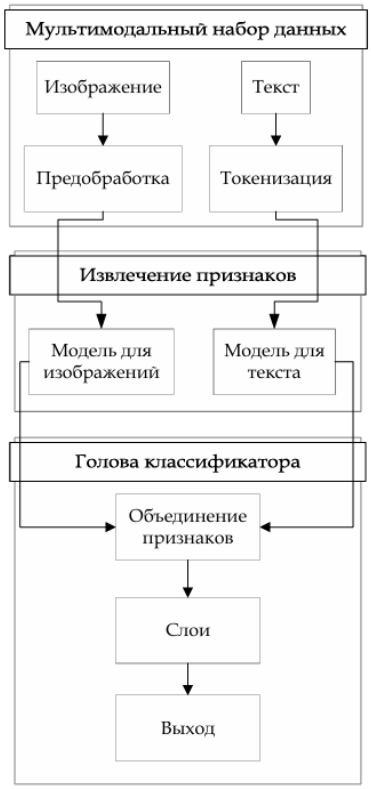


Рисунок 1 – Общая архитектура мультимодального классификатора

На рисунке 1 мы можем видеть, что после объединения текстовых и графических фичей они проходят через несколько слоев (где мы использовали нормализации и регуляризации чтобы избежать переобучения). Чтобы правильно сравнить все возможные модели, мы оценили 11 комбинаций (6 мультимодальных, а также 5 непосредственно унимодальных моделей). В качестве функции потерь автор выбрал бинарную кросс-энтропию с сигмоидой [18]:

$$l(x, y) = \frac{1}{N}\sum_{i=1}^{N} -w_i \times [y_i \times \log \sigma(x_i) + (1 - y_i) \times \log\big(1 - \sigma(x_i)\big)], \quad (1)$$

где x – входные логиты (перед применением сигмоиды), y – реальные метки, $\sigma(x_i)$ – сигмоида, N – размер батча, w_i – необязательный вес (для работы с несбалансированными данными). Что касается оптимизатора, то мы выбрали Adam (классический алгоритм Adam с регуляризацией веса) в паре с планировщиком скорости обучения Reduce on Plateau.

В качестве технической метрики мы выбрали ROC AUC с макро- и микро-усреднением. Первое обычно используется для оценки общего качества классификатора, не обращая внимания на дисбаланс в данных, однако второе учитывает его. В основном, в нашем

исследовании мы сосредоточились на обоих показателях, тем самым создав более стабильную модель как с общей, так и с точки зрения несбалансированности классов.

Результаты исследования

Унимодальные модели

Унимодальная часть представлена 5 моделями, из которых 3 текстовых и 2 для изображений. Основная идея заключается в сравнении полученных результатов между ними и дальнейшими мультимодальными версиями с точки зрения качества прогнозирования. Вообще говоря, текстовые модели легче и быстрее поддаются обучению и инференсу (в нашем случае в 8 раз быстрее обучение и в 5 раз инференс). Более того, их качество (с точки зрения ROC AUC) лучше, чем у классификаторов изображений. Полученные результаты представлены в таблице 2.

Таблица 2. Метрики качества унимодальных моделей

	Обучающая		**Валидационная**	
Унимодальные	**Макро**	**Микро**	**Макро**	**Микро**
Изображения				
Rdnet	83,7%	81,6%	75,8%	73,9%
Vit	81,8%	79,2%	74,0%	71,8%
Тексты				
E5	89,7%	86,0%	84,5%	79,2%
Sbert	90,9%	87,1%	86,0%	80,8%
Dpavlov	90,7%	87,1%	86,0%	80,9%

Как мы можем видеть из таблицы 2, текстовые модели, очевидно, работают лучше, чем модели для изображений. Также заметно, что текстовые менее переобучены, чем картиночные, поскольку абсолютная разница между всеми метриками (микро- и макро-) превышает 7%, а относительная - более 9,5%. В то же время, можно наблюдать и небольшое переобучение на микро-метриках и у текстовых классификаторов.

В целом, полученные результаты наглядно демонстрируют, что модели NLP (natural language preprocessing, обработка естественного языка) лучше, чем модели CV (computer vision, компьютерное зрение), поэтому, помимо проблемы переобучения текстовых классификаторов на микро-метриках, мы можем сделать вывод, что в бизнесе предпочтение следует отдавать текстовым моделям, а не картиночным в данном конкретном случае.

Мультимодальные модели

Мультимодальные модели представлены 6 комбинациями описанных ранее. Основная идея заключается в том, чтобы проверить, может ли комбинация текстовой и картиночной моделей улучшить общее качество классификатора или нет.

Таблица 3. Метрики качества мультимодальных моделей

	Обучающая		**Валидационная**	
Мультимодальные	**Макро**	**Микро**	**Макро**	**Микро**
Rdnet-E5	89,1%	85,8%	85,1%	81,2%
Rdnet-Sbert	89,3%	84,9%	85,4%	80,2%
Rdnet-Dpavlov	90,7%	86,2%	86,3%	81,0%
Vit-E5	89,3%	86,2%	84,5%	80,9%
Vit-Sbert	90,7%	88,9%	85,9%	82,8%
Vit-Dpavlov	89,4%	86,3%	85,1%	81,2%

Согласно данным таблицы 3, большинство мультимодальных моделей превосходят большинство унимодальных, однако только один мультимодальный классификатор (Rdnet-

Dpavlov) лучше, чем лидеры унимодальных (Sbert, Dpavlov). Более того, абсолютная и относительная разница между результатами обучения и валидации значительно ниже, чем в унимодальных моделях. При этом, ожидаемо, мультимодальные классификаторы более тяжелые и ресурсоемкие в обучении и инференсе, так как включают в себя картиночную модель.

Заключение

В данной работе автор провел полный цикл создания мультимодального классификатора. Был изучен опыт иностранных исследователей, затем были собраны и предобработаны данные из двух модальностей – синопсиса и постера, затем были обучены и оценены как унимодальные, так и мультимодальные модели и выбрана лучшая как с точки зрения метрик, так и с точки зрения переобучения. Наиболее эффективный классификатор был получен путем объединения картиночной модели Rdnet с текстовой моделью Dpavlov. Тем не менее, автор считает, что существуют дополнительные направления для исследований. Помимо использования альтернативных подходов к извлечению фичей, можно использовать дополнительные модальности и альтернативные модели машинного и глубокого обучения. Кроме того, эта работа может стать основой для дальнейших исследований в русскоязычном сегменте стриминговых сервисов.

Литература

1. Number of users in the video streaming (SVoD) segment of the digital media market Worldwide from 2017 to 2027. Statista. [Электронный ресурс] URL: https://www.statista.com/forecasts/1435540/number-of-users-video-streaming-svod-video-on-demand-market-worldwide (дата обращения: 10.02.2026)
2. Nam, J., & Tewfik, A. H. (n.d.). Combined audio and visual streams analysis for video sequence segmentation. 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing. doi:10.1109/icassp.1997.595337
3. Liu, Z., Wang, Y., & Chen, T. (1998). Audio Feature Extraction and Analysis for SceneSegmentation and Classification The Journal of VLSI Signal Processing, 20(1/2), 61–79.doi:10.1023/a:1008066223044
4. D. Brezeale and D. J. Cook, "Automatic Video Classification: A Survey of the Literature," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 38, no. 3, pp. 416-430, May 2008, doi: 10.1109/TSMCC.2008.919173.
5. Jain, S. K., & Jadon, R. S. (2009). Movies genres classifier using neural network. 2009 24th International Symposium on Computer and Information Sciences. doi:10.1109/iscis.2009.5291884
6. Howard Zhou, Tucker Hermans, Asmita V Karandikar, and James M Rehg. Movie genre classification via scene categorization. In Proceedings of the 18th ACM international conference on Multimedia, pages 747–750. ACM, 2010.
7. Giuseppe Portolese and Valéria Delisandra Feltrim. On the use of synopsis-based features for film genre classification. In Anais do XV Encontro Nacional de Inteligência Artificial e Computacional, pages 892–902. SBC, 2018.
8. Paola Cascante-Bonilla, Kalpathy Sitaraman, Mengjia Luo, and Vicente Ordonez. Moviescope: Large-scale analysis of movies using multiple modalities. arXiv preprint arXiv:1908.03180, 2019.
9. Jônatas Wehrmann and Rodrigo C Barros. Movie genre classification: A multi-label approach based on convolutions through time. Applied Soft Computing, 61:973–982, 2017.
10. Mangolin, R.B., Pereira, R.M., Britto, A.S. et al. A multimodal approach for multi-label movie genre classification. Multimed Tools Appl 81, 19071–19096 (2022). https://doi.org/10.1007/s11042-020-10086-2
11. Andriyanov, N.A. Combining Text and Image Analysis Methods for Solving Multimodal Classification Problems. Pattern Recognit. Image Anal. 32, 489–494 (2022). https://doi.org/10.1134/S1054661822030026

12. Портал Кинопоиск [Электронный ресурс] URL: https://www.kinopoisk.ru/ (дата обращения: 10.02.2026)
13. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv. [Электронный ресурс] URL: https://arxiv.org/abs/2010.11929 (дата обращения: 10.02.2026)
14. Kim, D., Heo, B., & Han, D. (2024). DenseNets reloaded: Paradigm shift beyond ResNets and ViTs. arXiv. [Электронный ресурс] URL: https://arxiv.org/abs/2403.19588 (дата обращения: 10.02.2026)
15. SberDevices. (2021, June 10). Модель ruGPT-3 от Сбер: особенности, различия с GPT-3, примеры генерации. Habr. [Электронный ресурс] URL: https://habr.com/ru/companies/sberdevices/articles/560748/ (дата обращения: 10.02.2026)
16. Kuratov, Y., & Arkhipov, M. (2019). Adaptation of deep bidirectional multilingual transformers for Russian language. arXiv. [Электронный ресурс] URL: https://doi.org/10.48550/arXiv.1905.07213 (дата обращения: 10.02.2026)
17. Wang, L., Yang, N., Huang, X., Yang, L., Majumder, R., & Wei, F. (2024). Multilingual E5 text embeddings: A technical report. arXiv. [Электронный ресурс] https://arxiv.org/abs/2402.05672 (дата обращения: 10.02.2026)
18. PyTorch. (n.d.). torch.nn.BCEWithLogitsLoss. PyTorch. [Электронный ресурс] URL: https://pytorch.org/docs/stable/generated/torch.nn.BCEWithLogitsLoss.html (дата обращения: 10.02.2026)

THE USE OF MULTIMODAL MODELS IN DETERMINING FILM GENRES

Shuvarikov M.D.

Financial University under the Government of the Russian Federation

The use of various data modalities makes it possible to successfully improve the quality of models. One of the areas where business interest in using this approach is growing is the prediction of movie genres on streaming services, the number of users of which has been increasing especially rapidly in recent years. The purpose of this work is to explore existing approaches to predicting film genres and propose our best multimodal model trained on data from the Russian Kinopoisk service.

ИСПОЛЬЗОВАНИЕ МЕТОДА ТЕСТОВЫХ ИСПЫТАНИЙ ДЛЯ РЕШЕНИЯ ПРАКТИЧЕСКИХ ЗАДАЧ РЕКОНСТРУКЦИИ ОДНОМЕРНЫХ СИГНАЛОВ, ИСКАЖЕННЫХ СИСТЕМАМИ ОБРАБОТКИ

с.н.с., к.т.н. Новиков-Бородин А.В.
Институт ядерных исследований РАН, 117312 Москва, Россия

Аннотация: Рассматриваются возможности применения математических методов тестовых испытаний для решения практических задач реконструкции сигналов, искаженных при обработке линейными стационарными системами, описываемыми уравнениями типа свертки. Методы тестовых испытаний позволяют осуществить неслепую реконструкцию сигналов без определения передаточных функций систем обработки вместо которых используются тестовые сигналы этих систем. Использование тестовых сигналов из класса основных функций вместо передаточных функций из класса обобщенных значительно упрощает процедуру реконструкции и делает ее более точной и устойчивой. Операторный подход, основанный на уравнении многомерной свертки, повышает скорость численных вычислений. Для решения некорректно поставленных и плохо обусловленных задач используется техника регуляризации, позволяющая эффективно восстанавливать реальные недетерминированные сигналы с высоким уровнем шумов и неопределенностей. Приводятся примеры практического использования методов, процедура реконструкции сравнивается с существующими подходами. Предлагаемая методика позволяет расширить область применения математических методов реконструкции, повысить качество обработки сигналов и изображений без использования сложных и дорогостоящих аппаратных методов усовершенствования оборудования.
Программные коды и наборы данных доступны по адресу: https://github.com/novikov-borodin/data-rec.
Ключевые слова: математическая обработка сигналов; линейные стационарные системы; тестовые испытания систем обработки; уравнение типа свертки; некорректно поставленные и плохо обусловленные задачи; деконволюция; техника регуляризации.

Введение

Реальные системы обработки несовершенны, что приводит к системным искажениям и появлению случайных шумов в обрабатываемых ими сигналах. Аппаратные методы уменьшения искажений связаны с решением сложных научно-технических задач и созданием дорогостоящего оборудования, а альтернативные математические методы сталкиваются с серьезными трудностями при решении некорректно поставленных и плохо обусловленных задач реконструкции, как правило, встающих на практике (см., например, [1]).

Линейные стационарные (ЛС) системы, рассматриваемые в данной работе, удовлетворяют уравнению типа свертки (см. раздел 1) и описываются передаточными функциями, характеризующими системные искажения при измерении, приеме, передаче или другой обработке сигналов. В неслепых методах математической реконструкции сигналов используются передаточные функции систем обработки, но эти функции принадлежат к классу обобщенных, и их определение является достаточно сложной задачей, что резко ограничивает точность неслепых методов и область их применения некоторыми частными случаями известных видов искажений и/или шумов (см., например, [2-4]). Методы слепого восстановления (см., например, [5]) позволяют избежать проблем с определением передаточных функций, но они основаны на априорных знаниях или предположениях об исходном или желаемом виде реконструируемых данных. Реальные сигналы, искажения и шумы, вносимые системами обработки, очень разнообразны, что сильно ограничивает область применения слепых методов. Обойтись без прямого определения передаточных функций позволяют нейронные сети (см., например, [6]), но их обучение происходит на большом количестве примеров, что дает усредненные результаты и во многих случаях не обеспечивает высокой точности реконструкции. По общим оценкам, приведенным в пленарном докладе [7], существующие математические методы, включая нейронные сети, позволяют повысить эффективные параметры измерительных систем (например, эффективную полосу пропускания) в среднем в $1{,}5 - 2$ раза.

В данной работе рассматриваются особенности и примеры практического применения метода тестовых испытаний (ТИ) [8,9], совмещающего в себе отдельные черты перечисленных методов, определяющие его преимущества. Формально, ТИ-метод не использует передаточные функции систем обработки и не является неслепым, но ТИ-уравнение, лежащее в его основе, устанавливает точное соответствие между сигналами. Использование тестовых сигналов вместо передаточных функций можно рассматривать как процесс обучения нейросети при поиске обратного преобразования. В отличие от передаточных функций, тестовые сигналы принадлежат

к классу основных функций, поэтому их определение, как численными, так и экспериментальными методами, не вызывает больших проблем, что определяет простоту, универсальность и точность предлагаемого подхода. При отсутствии передаточных функций из класса обобщенных появляется возможность эффективного использования существующих методов решения некорректно поставленных и плохо обусловленных задач реконструкции, таких как методы регуляризации (см., например, [12]), используемые в данной работе, или итерационные методы, вейвлет анализ и др. Эти особенности обеспечивают универсальность предлагаемого подхода, точность восстановления и устойчивость алгоритма даже при высоком уровне шумов и неопределенностей в данных. В работе приводятся примеры реконструкции сигналов, проводится сравнение предлагаемой методики с существующими подходами.

1. Метод тестовых испытаний

Функциональная система обработки сигналов называется линейной стационарной, если в области $x \in D$ она характеризуется передаточной функцией $u(x)$ и ее реакция $w(x)$ (выходной сигнал) на внешнее воздействие $v(x)$ (входной сигнал) описывается уравнением типа свертки, которое будем называть уравнением ЛС-систем, или, кратко, ЛС-уравнением:

$$w(x) - \{\eta(x)\} = (u * v)(x) = \int_D u(x-\xi)v(\xi)\mathrm{d}\xi. \qquad (1)$$

Неопределенность ЛС-уравнения: $\eta(x) = (\Delta w - u * \Delta v - \Delta u * v + \Delta u * \Delta v)(x)$ можно определить из уравнения свертки: $w_0(x) = (u_0 * v_0)(x)$ детерминированных функций без неопределенностей: $w_0(x) = (w - \Delta w)(x)$, $u_0(x) = (u - \Delta u)(x)$ и $v_0(x) = (v - \Delta v)(x)$.

Вводя оператор свертки $\mathcal{F}$ с порождающей функцией $f(x)$: $\mathcal{F}g(x) = (f * g)(x)$, ЛС-уравнение можно представить в операторной форме: $\mathcal{U}v(x) = w(x) - \{\eta(x)\}$, а используя теорему Бореля о свертке, – в спектральной форме: $U(\omega)V(\omega) = W(\omega) - \{N(\omega)\}$, где $F(\omega) = FT[\mathcal{F}] = FT[f(x)]$ – спектр (преобразование Фурье) функции $f(x)$ и оператора $\mathcal{F}$, и $N(\omega) = \Delta W(\omega) - U(\omega)\Delta V(\omega) - \Delta U(\omega)V(\omega) - \Delta U(\omega)\Delta V(\omega)$. Спектральная форма ЛС-уравнения имеет простой вид, но для определения спектров функции должны быть известны во всем пространстве $\mathbb{R}$, а не только в заданной области D, что резко ограничивает область применения спектральных методов.

ЛС-уравнение в дискретном виде представляет собой частичную сумму интеграла (1) на сетке $\{d_i\}, i = 1 \dots m,$ разбивающей область D:

$$w_i - \{\eta_i\} = \sum_{j=1}^{m} u_j \; v_{i-j} \, \Delta_j \, . \qquad (2)$$

Здесь $w = \{w_i\}$, $\eta = \{\eta_i\}$, $u = \{u_i\}$ и $v = \{v_i\}$ – векторы, соответствующие функциям $w(x)$, $\eta(x)$, $u(x)$ и $v(x)$, Δ_j – элемент объема. Если переопределить u_j как $u_j\Delta_\mathrm{j}$ или на равномерной сетке разбиения положить $\Delta_j = 1$, то в правой части уравнения (2) получим выражение дискретной свертки векторов $u = \{u_i\}$ и $v = \{v_i\}$.

В общем случае, передаточные функции ЛИ-систем принадлежат к классу обобщенных, что, с учетом, как правило, некорректно поставленных задач реконструкции и нестабильности решений, создает серьезные трудности как при их определении, так и при использовании в методах математической реконструкции сигналов. Для решения этих проблем в работах [8,9] был предложен метод тестовых испытаний, в основе которого лежит уравнение тестовых испытаний (ТИ-уравнение):

$$(v_c * w)(x) - (w_c * v)(x) = \{\eta_T(x)\}, \qquad (3)$$

получающееся из ЛС-уравнений (1) для обрабатываемых сигналов $v(x)$, $w(x)$ и тестовых сигналов $v_c(x)$, $w_c(x)$ при вычитании их сверток с функциями $v_c(x)$ и $v(x)$ и исключении передаточной функции. При этом неопределенность ТИ-уравнения (3): $\eta_T(x) = \big(v_c * (\Delta w - u_0 * \Delta v) - v * (\Delta w_c - u_0 * \Delta v_c)\big)(x)$ оказывается меньше, чем неопределенности $\eta(x)$ и $\eta_c(x)$ исходных ЛИ-уравнений за счет сокращения членов, входящих в $\eta(x)$ и $\eta_c(x)$ с противоположными знаками. Кроме того, ТИ-уравнение не содержит передаточных функций и при его использовании не возникает ошибок представления функций из класса обобщенных. Все

это приводит к снижению ожидаемой погрешности реконструкции сигналов из ТИ-уравнения и повышению устойчивости алгоритма восстановления.

В спектральной форме ТИ-уравнение имеет вид: $W_c(\omega)V(\omega) = V_c(\omega)W(\omega) - \{N_T(\omega)\}$, а в операторной форме: $\mathcal{W}_c v(x) = \mathcal{V}_c w(x) - \{\eta_T(x)\}$, где $\mathcal{W}_c$ и $\mathcal{V}_c$ – операторы свертки с производящими функциями $w_c(x)$ и $v_c(x)$. В дискретном виде ТИ-уравнения: $\mathcal{V}_c w = \mathcal{W}_c v - \{\eta_T\}$ многомерные операторы свертки понимаются в смысле (2). На равномерной сетке разбиения элементы объема в обеих частях этого уравнения сокращаются, и в отличие от уравнения (2), переопределение функций здесь не требуется.

Сигналы ЛС-систем относятся к функциям основного класса, и их определение не вызывает затруднений, поскольку могут быть использованы как численные, так и экспериментальные методы. При экспериментальном определении тестовых сигналов не требуется ни анализ ЛИ-систем, ни учет паразитных параметров их элементов, так как эта информация содержится в тестовых сигналах. Это упрощает процедуру реконструкции и позволяет восстанавливать сигналы высокоскоростных и прецизионных систем обработки практически любой сложности.

2. Решение ТИ-уравнений

ТИ-уравнение является уравнением типа свертки. При этом оно не содержит передаточных функций из класса обобщенных, а лишь сигналы из класса основных функций, что позволяет эффективно применять существующие методы решения некорректно поставленных и плохо обусловленных задач реконструкции сигналов с высоким уровнем шумов и неопределенностей, такие, например, как техника регуляризации, итерационные методы, вейвлет анализ и т.д. В данной работе будет использоваться техника регуляризации, которая не является интегральной частью ТИ-метода, т.к. и другие методы можно можно применять для решения ТИ-уравнений.

Согласно теории регуляризации (см., напр., [10]), решением $v_r(x)$ ТИ-уравнения будет функция, минимизирующая функционал $\Phi[v] = \|\mathcal{V}_c w - \mathcal{W}_c v\|^2 + \lambda\|\mathcal{P}v\|^2$ со стабилизатором $\lambda\|\mathcal{P}v\|^2$, где $\mathcal{P}$ – оператор ограничений, накладываемых на реконструируемую функцию $v(x)$, а λ – параметр регуляризации. Для реконструкции сигналов наиболее широкого класса, мы не будем накладывать дополнительные ограничения $\mathcal{P}$ на решения ТИ-уравнений, полагая: $\mathcal{P}^+\mathcal{P} = \mathcal{I}$, где $\mathcal{I}$ – тождественный оператор, что соответствует поиску решений со “степенью непрерывности” определяемой значением λ. Приравнивая к нулю вариацию функционала $\Phi[v]$ по v, получим:

$$v_r(x) = (\mathcal{W}_c^+\mathcal{W}_c + \lambda\mathcal{I})^{-1}\mathcal{W}_c^+\mathcal{V}_c w(x). \qquad (4)$$

Здесь $\mathcal{W}_c^+$ – сопряженно-транспонированный оператор к оператору $\mathcal{W}_c$ с производящей функцией $w_c(x)$. Остаток $\left\|\left(v_c * w - w_c * v_r(\lambda)\right)(x)\right\|$ монотонно возрастает с увеличением λ, поэтому оптимальное значение λ можно найти из уравнения: $\left\|\left(v_c * w - w_c * v_r(\lambda)\right)(x)\right\| = \|\eta_T(x)\|$. Отсюда следует и необходимое условие для восстановления сигналов из ТИ-уравнения: $\|\eta_T(x)\| \le \|(v_c * w - w_c * v)(x)\|$. В спектральной форме решение (4) имеет вид: $V_r(\omega) = W_c^+(\omega)V_c(\omega)W(\omega)/(W_c^+(\omega)W_c(\omega) + \lambda)$.

Из (4) получим дискретное регуляризированное решение ТИ-уравнения в операторном:

$$v_r = (\mathcal{W}_c^+\mathcal{W}_c + \lambda\mathcal{I})^{-1}\mathcal{W}_c^+\mathcal{V}_c w \qquad (5)$$

и в спектральном виде: $V^r_{j_1..j_M} = W^{c+}_{j_1..j_M}V^c_{j_1..j_M}W_{j_1..j_M}/\left(W^{c+}_{j_1..j_M}W^c_{j_1..j_M} + \lambda\right)$.

Количество операций необходимых для восстановления массивов с n элементами с помощью операторного уравнения (5), основанного на уравнении дискретной свертки (2), может быть оценено как: $N(n) \sim O(5n^2 + n + 1)$, что для больших n намного быстрее, чем, например, вычисления свертки одномерных сигналов с помощью матриц Тёплица, требующих $\sim 5n^3$ операций. Необходимое количество операций при использовании спектральной формы уравнения (5) (с прямым и обратным ДПФ) равно: $N(n) \sim O(4n\log(n) + n + 1)$.

Спектральные методы (5) быстрее операторных, но возможности их применения ограничены, поскольку для определения спектров необходимо знать значения функций во всем пространстве, а сигналы, зачастую, известны лишь на ограниченных областях D, за пределами которых их значения могут быть не равны нулю.

3. Примеры реконструкции и обсуждение

Результаты численного эксперимента по восстановлению одномерных сигналов сложной формы (суперпозиция трапецеидального и гауссового импульсов) на входе емкостного делителя напряжения из искаженного выходного сигнала $w(t)$ по тестовым сигналам $w_c(t)$ и $v_c(t)$ представлены на рисунке 1.

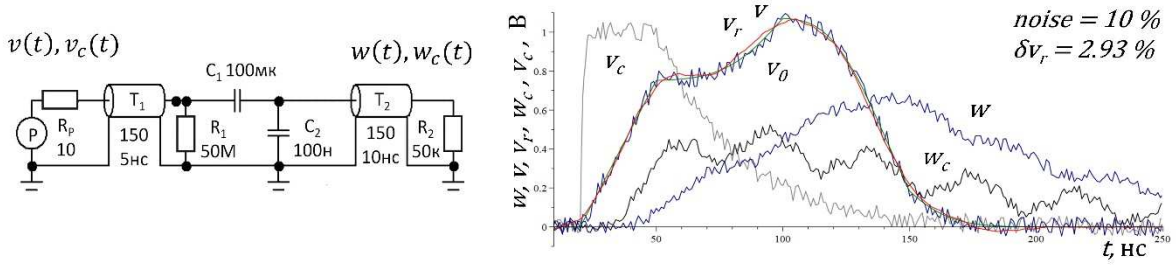


Рисунок 1 – Восстановление динамических сигналов

В качестве тестового воздействия $v_c(t)$ использовался трапецеидальный импульс с разными экспоненциальными фронтами, обеспечивающий широкий спектр сигнала (подробности см. в работах [8,9]). Исходные сигналы были рассчитаны с помощью симулятора аналоговых/цифровых схем Micro-CAP 12.2.0.4 компании Spectrum Software. Принципиальная схема делителя была взята из стандартной библиотеки этой программы. Частота среза делителя составляла 1 МГц, но с помощью математической реконструкции ТИ-методом удалось восстановить наносекундные сигналы с верхней границей спектра больше 1 ГГц. Таким образом, эффективная полоса пропускания системы обработки была улучшена на три порядка. Относительные ошибки реконструкции δv_r при наличии гауссовского шума с дисперсией 5 и 10 % от максимальных значений сигналов v, v_c и w, w_c составили 1.73 и 2.93 %. Численные расчеты проводились в среде Maple 2017.3 Waterloo Maple Inc. на ноутбуке HP 255 G7 с процессором AMD Ryzen 3, 2.5 ГГц. Время восстановления дискретных сигналов размером $n = 501$, соответствующих шагу дискретизации $\Delta = 0.5$ нс составило 1.6..2.0 сек. При размере векторов $n = 251$ ($\Delta = 1.0$ нс) время вычислений составило 0.4…0.5 сек, что соответствует сделанным ранее оценкам квадратичной зависимости времени вычислений от размера данных (см. комментарии к уравнению (5)). Коды программ доступны в библиотеке GitHub по адресу https://github.com/novikov-borodin/data-rec/signal-rec. При использовании программ ссылки на работы автора и директории автора в библиотеке GitHub обязательны.

Рост ошибок реконструкции нелинейно замедляется с ростом уровня шумов в данных, что объясняется эффективным сглаживанием амплитуды спектра высокочастотного шума при регуляризации решения. В общем случае ошибка реконструкции является случайной, поскольку зависит не только от уровня шумов, но и от их распределений в данных. На рисунке 2 представлен график зависимости относительной ошибки реконструкции $\delta \boldsymbol{v}_r$, % от относительного уровня шума $\delta \boldsymbol{w}$, % в выходном сигнале.

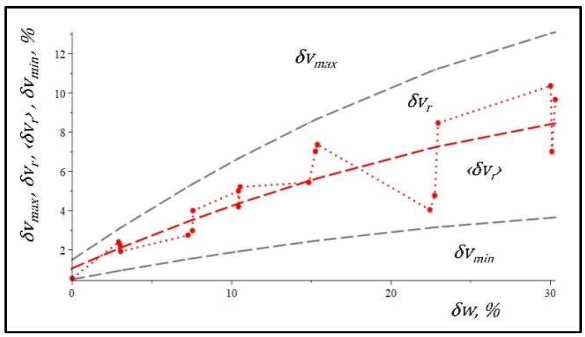


Рисунок 2 – Зависимость погрешности реконструкции от уровня шумов в сигналах

Точками показаны ошибки реконструкции, полученные в численных экспериментах. Кривая $\langle \delta \boldsymbol{v}_r \rangle$ соответствует ожидаемым значениям ошибок реконструкции, а кривые $\delta \boldsymbol{v}_{min}$ и $\delta \boldsymbol{v}_{max}$

показывают ожидаемый разброс этих ошибок. Уровень шумов в тестовых сигналах на рисунке 2 составлял $\delta \boldsymbol{w}_c = 2.2$ % и $\delta \boldsymbol{v}_c = 1.4$ %.

В работе [11] математические методы использовались для расширения полосы пропускания быстродействующего осциллографа путем компенсации его частотной характеристики. Для этого на вход осциллографа подавался тестовый сигнал, аппроксимирующий дельта-функцию, и спектр отклика, приближенно совпадающий с импульсной характеристикой осциллографа (его передаточной функцией), использовался для корректировки спектральной характеристики осциллографа и расширения его эффективной полосы пропускания. Такой подход кроме погрешностей измерений включает в себя как погрешности аппроксимации дельта-функции входным сигналом и импульсной характеристики осциллографа соответствующим выходным сигналом, так и ограничения, связанные с использованием спектрального подхода к восстановлению сигналов.

В работе [12] была предпринята попытка уменьшить погрешность аппроксимации дельта-функции входным сигналом. Для этого на вход осциллографа подавался сигнал в виде ступенчатой функции $\Theta(t)$ с последующим дифференцированием отклика осциллографа для получения его импульсной характеристики. Действительно, согласно свойствам свертки из (1) следует: $w'(t) - \{\eta'(t)\} = (u * \Theta')(t) = (u * \delta)(t) = u(t)$. Такой подход сопряжен с погрешностями формирования входного сигнала в виде ступеньки и с резким ростом уровня высокочастотных шумов и неопределенности уравнения (1) при дифференцировании, поскольку согласно преобразованию Фурье: $FT[\eta'(t)] = i\omega\, FT[\eta(t)]$. Эксперимент в работе [12] показал, что при увеличении эффективной полосы пропускания осциллографа в 1,5 раза среднеквадратичное значение шума в восстановленном импульсном сигнале увеличивается в 1,5–3 раза. Отметим, что при таком подходе также сохраняются все трудности и ограничения, связанные с использованием спектрального подхода к восстановлению сигналов.

Метод тестовых испытаний позволяет избежать многих проблем, подобных [11,12]. Действительно, в ТИ-методах входной и выходной тестовые сигналы осциллографа могут быть непосредственно использованы для реконструкции сигналов без определения импульсной характеристики или его спектра. При этом, в ТИ-методах не требуется ни спектральных преобразований, приводящих к ограничениям спектрального подхода, ни дополнительных операций, таких как дифференцирование сигналов, приводящих к увеличению уровня шумов и, следовательно, погрешности реконструкции.

Заключение

Рассмотрены возможности применения на практике математических методов восстановления сигналов, искаженных и зашумленных при их обработке линейными стационарными системами, с помощью тестовых испытаний этих систем. В предлагаемых методах реконструкции вместо передаточных функций систем обработки, являющихся также функциями искажения сигналов, используются тестовые сигналы систем, что можно назвать динамической калибровкой систем обработки сигналов. Использование тестовых сигналов значительно упрощает процедуру восстановления и делает ее более устойчивой и точной. Операторный подход, основанный на уравнении свертки, существенно повышает скорость численных расчетов. Предложенная методика расширяет возможности применения методов математической реконструкции на практике, позволяет повысить качество обработки сигналов и изображений без сложной и дорогостоящей модернизации оборудования. Методика имеет широкую область практического применения.

Литература

1. Kazufumi I., Bangti J. Inverse Problems: Tikhonov Theory and Algorithms // Singapore: World Scientific, 2014.
2. Chaganova O.B., Grigoryev A.S., Nikolaev D.P., Nikolaev I.P. Applied aspects of modern non-blind image deconvolution methods // Computer Optics, 48(4), 2024. doi: 10.18287/2412-6179-CO-1409.

3. Wang H. et al. Ambulatory ECG noise reduction algorithm for conditional diffusion model based on multi-kernel convolutional transformer // The Review of scientific instruments. 95(9), 2024.
4. Junjie T. A.O., et al. An image defocus deblurring method based on gradient difference of boundary neighborhood // Virtual Reality & Intelligent Hardware, 5(6): 538-549, 2023. doi: 10.1016/j.vrih.2023.06.008.
5. Levin A., et al. Understanding and evaluating blind deconvolution algorithms // 2009 IEEE conference on computer vision and pattern recognition, IEEE, 2009.
6. Hyeontae Jo, Hwijae Son, Hyung Ju Hwang, Eun Heui Kim. (2020). Deep neural network approach to forward-inverse problems. // Networks and Heterogeneous Media, 15(2): 247-259. doi: 10.3934/nhm.2020011.
7. Malay I. Digital Signal Processing in Modern Radio Engineering Measurements // Plenary Talk at XXV International Conference Digital Signal Processing and Its Applications – DSPA-2023, 2023.
8. Novikov-Borodin A.V. Using Test Trials of Processing Systems for Non-Blind Recovery of Signals and Images // Journal of Data Science and Intelligent Systems, 2025. doi: 10.47852/bonviewJDSIS52024990.Новиков-Бородин А.В. Математическая реконструкция сигналов и изображений с помощью тестовых испытаний: неслепой подход. Журнал вычислительной математики и математической физики, т.65, № 3, 2025, с.401-414. doi: 10.31857/S0044466925030135.Elsayed M.A., El-Shafai W., Rashwan M.A., Dessouky M.I., El-Fishawy A.S., Abd El-Samie F.E. Efficient iterative implementation of regularized solutions for image and signal reconstruction problems // Journal of Optics, 2023. doi: 10.1007/s12596-023-01179-w.
11. Zhu J., Wang Y., Zhao K., Wang Y. Oscilloscope Measurement Bandwidth Expansion Using Frequency Response Compensation // IEEE Access, vol. 12: 49772-49778, 2024. doi: 10.1109/ACCESS.2024.3384381.
12. Kleopin A.V., Zenchenko M.A. Tikhonov Regularization in Pulse Signal Processing for Oscilloscope Measurements // XXIV International Conference on Digital Signal Processing and its Applications DSPA-2022, 2022. doi: 10.1109/DSPA53304.2022.9790777.

APPLICATION OF TEST TRIAL METHODS FOR SOLVING THE PRACTICAL TASKS OF RECOVERY OF ONE-DIMENSIONAL SIGNALS DISTORTED BY PROCESSING SYSTEMS

A.V. Novikov-Borodin
Institute for Nuclear Research of RAS, 117312 Moscow, Russia

Abstract. This article explores the potential of using the mathematical methods of test trials to solve practical problems of reconstructing signals, distorted during processing by linear stationary (time-invariant) systems described by a convolution-type equations. The test trial methods enable non-blind signal reconstruction without determining the transfer functions of processing systems. The use of test signals belonging to the class of ordinary functions instead of the transfer functions from the generalized functions significantly simplifies the reconstruction procedure, making it more accurate and robust. An operator approach based on the convolution equation increases the speed of numerical computations. To solve ill-posed and ill-conditioned problems, a regularization technique is used, allowing for the efficient reconstruction of non-deterministic signals with high levels of noise and uncertainty. Examples of the methods' application are provided, and the reconstruction procedure is compared with existing approaches. The proposed method allows us to expand the scope of application of mathematical reconstruction methods and improve the quality of signal and image processing without the use of complex and expensive hardware modification methods.
Program codes and datasets are openly available in GitHub library at: https://github.com/novikov-borodin/data-rec.
Keywords: mathematical signal processing; linear stationary (time-invariant) systems; test trials of processing systems; convolution type equation; ill-posed and ill-conditioned problems; deconvolution; regularization technique.

ПРИМЕНЕНИЕ МЕТОДА ПРЯМЫХ СДВИГОВ ДЛЯ РЕШЕНИЯ ПРАКТИЧЕСКИХ ЗАДАЧ ОБРАБОТКИ ОДНОМЕРНЫХ СИГНАЛОВ

с.н.с., к.т.н. Новиков-Бородин А.В.
Институт ядерных исследований РАН, 117312 Москва, Россия

Аннотация: Рассматриваются возможности применения математических методов прямых сдвигов для коррекции, реконструкции и моделирования сигналов, искаженных линейными стационарными системами в процессе обработки. Линейные стационарные системы охватывают широкий класс систем обработки, описываемых уравнениями типа свертки. Методы прямых сдвигов позволяют в ряде важных и часто встречающихся на практике случаях решать задачи обработки сигналов при минимальной исходной информации с минимальными вычислительными ресурсами и погрешностями. Эффективность метода прямых сдвигов обусловлена возможностью выбора вида преобразований и сокращения их числа. Рассмотрены примеры использования метода прямых сдвигов на практике. Приведены результаты сравнения метода прямых сдвигов с методами Гаусса-Жордана, прямой матричной инверсии, LU- и SVD-разложений.

Программные коды доступны в библиотеке GitHub по адресу: https://github.com/novikov-borodin/lst-data-proc.

Ключевые слова: математическая обработка сигналов; метод прямых сдвигов; линейные стационарные системы; уравнение типа свертки.

Введение

Несовершенство систем обработки приводит к появлению искажений в получаемых, преобразуемых и передаваемых ими данных. Аппаратные методы уменьшения искажений связаны с решением сложных научно-технических и технологических задач, а также с созданием дорогостоящего оборудования. Альтернативой им являются математические методы реконструкции неискаженных данных из искаженных в процессе обработки (см., например, [1,2] и ссылки в них). Предлагаемые в данной работе математические методы прямых сдвигов (ПС) предназначены для коррекции, реконструкции и моделирования сигналов и изображений, искаженных в процессе их обработки линейными стационарными (ЛС) системами. ЛС-системы охватывают широкий класс систем обработки, описываемых уравнениями типа свертки (см., например, [3]).

Методы прямых сдвигов позволяют решать задачи обработки сигналов при минимуме исходной информации с минимальными погрешностями и вычислительными ресурсами. Эффективность ПС-метода обусловлена алгоритмом его реализации с возможностью выбора вида преобразований и сокращения их числа. В работе рассмотрены примеры использования прямых сдвигов для решения часто встречающихся на практике типовых задач. Приводятся результаты сравнения ПС-метода с другими методами прямой инверсии, такими как метод Гаусса-Жордана, прямой матричной инверсии, LU- и SVD- разложений.

1. Метод прямых сдвигов

Функциональную систему обработки одномерных или многомерных сигналов, включая изображения, будем называть линейной стационарной, если в области $x \in D \subset \mathbb{R}$ она характеризуется передаточной функцией $u(x)$ и ее реакция $w(x)$ (выходной сигнал) на внешнее воздействие $v(x)$ (входной сигнал) описывается уравнением типа свертки, которое будем называть уравнением ЛС-систем, или, кратко, ЛС-уравнением:

$$w(x) - \{\eta(x)\} = (u * v)(x) = \int_D u(x-\xi)v(\xi)\mathrm{d}\xi. \qquad (1)$$

Неопределенность ЛС-уравнения: $\eta(x) = (\Delta w - u * \Delta v - \Delta u * v + \Delta u * \Delta v)(x)$ можно определить из уравнения свертки $w_0(x) = (u_0 * v_0)(x)$ детерминированных функций без неопределенностей: $w_0(x) = (w - \Delta w)(x)$, $u_0(x) = (u - \Delta u)(x)$ и $v_0(x) = (v - \Delta v)(x)$.

Вводя оператор свертки $\mathcal{F}$ с порождающей функцией $f(x)$: $\mathcal{F}g(x) = (f * g)(x)$, ЛС-уравнение можно представить в операторной форме: $\mathcal{U}v(x) = w(x) - \{\eta(x)\}$, а используя теорему Бореля о свертке, – в спектральной форме: $U(\omega)V(\omega) = W(\omega) - \{N(\omega)\}$, где $F(\omega) = FT[\mathcal{F}] = FT[f(x)]$ – спектр (преобразование Фурье) функции $f(x)$ и оператора $\mathcal{F}$, и $N(\omega) = \Delta W(\omega) - U(\omega)\Delta V(\omega) - \Delta U(\omega)V(\omega) - \Delta U(\omega)\Delta V(\omega)$. Спектральная форма ЛС-уравнения имеет простой вид, но для определения спектров функции должны быть известны во всем пространстве

$\mathbb{R}$, а не только в заданной области D, что резко ограничивает область применения спектральных методов.

ЛС-уравнение в дискретном виде представляет собой частичную сумму интеграла (1) на сетке $\{d_i\}, i = 1 \dots m$, разбивающей область D:

$$w_i - \{\eta_i\} = \sum_{j=1}^{m} u_j \; v_{i-j} \, \Delta_j \, . \qquad (2)$$

Здесь $w = \{w_i\}$, $\eta = \{\eta_i\}$, $u = \{u_i\}$ и $v = \{v_i\}$ – векторы, соответствующие функциям $w(x)$, $\eta(x)$, $u(x)$ и $v(x)$, Δ_j – элемент объема. Если переопределить u_j как $u_j \Delta_j$ или на равномерной сетке разбиения положить $\Delta_j = 1$, то в правой части уравнения (2) получим выражение дискретной свертки векторов $u = \{u_i\}$ и $v = \{v_i\}$.

Сдвиг функции $f(x)$ можно осуществить с помощью оператора свертки $\mathcal{S}$ с порождающей функцией $s(x)$: $f(x - x_0) = \mathcal{S}f(x) = (s * f)(x)$, где $s(x) = \delta(x - x_0)$ и $\delta(x)$ – дельта-функция Дирака. Согласно линейности, коммутативности и ассоциативности свертки последовательное применение $\mathcal{S} = \mathcal{S}_n \dots \mathcal{S}_2 \mathcal{S}_1$ и линейные комбинации $\mathcal{S} = \sum_i a_i \mathcal{S}_i$ операторов сдвига являются операторами свертки. Осуществляя преобразования $\mathcal{S}$ над обеими частями ЛС-уравнения, получим уравнение прямых сдвигов (ПС):

$$\mathcal{S}w(x) - \{\mathcal{S}\eta(x)\} = (s * u * v)(x) = (\mathcal{S}u * v)(x) = (u * \mathcal{S}v)(x). \qquad (3)$$

Таким образом, осуществляя ПС-преобразования над обеими частями ЛС-уравнения можно из отклика $w(x)$ ЛС-системы на входной сигнал $v(x)$ смоделировать ее отклик $\mathcal{S}w(x)$ на входной сигнал $\mathcal{S}v(x)$, не определяя ее передаточную функцию; или же отклик $\mathcal{S}w(x)$ системы с передаточной функцией $\mathcal{S}u(x)$ на такой же входной сигнал, не определяя его. Согласно (3) ошибка преобразований составит $\{\mathcal{S}\eta(x)\}$.

2. Базовая задача ПС-преобразований

Эффективность ПС-методов проявляется при решении типовых задач, сводимых к базовой задаче о наложении двух импульсов.

Задача (наложение двух импульсов). Пусть известен отклик $w(x)$ ЛС-системы (1) от двух смещенных относительно друг друга на расстояние (или время) $T > 0$ импульсов $v_0(x)$: $v(x) = v_0(x) - a v_0(x - T)$ ($w(x) = v(x) = 0$ при $x < 0$). Реконструировать на интервале $x \in (-\infty, L)$, $L > T$, отклик $w_0(x)$ этой системы от импульса $v_0(x)$.

Из ЛС-уравнения (1) имеем: $w(x) = (u * s * v_0)(x) = (s * w_0)(x), s(x) = \delta(x) - a\delta(x - T)$. Применяя к функции $s(x)$ последовательность ПС-преобразований: $\mathcal{S} = \mathcal{S}_n \cdots \mathcal{S}_1$, где $\mathcal{S}_k$: $\delta(x) + a^{2^{k-1}} w(x - 2^{k-1}T)$, получим: $\mathcal{S}s(x) = \delta(x) - a^{2^n}\delta(x - 2^n T)$. Тогда, согласно (2): $\mathcal{S}w(x) = (\mathcal{S}s * w_0)(x) = w_0(x) - a^{2^n} w_0(x - 2^n T)$. При $2^n T > L$ на интервале $x \in (-\infty, L)$ имеем: $\mathcal{S}w(x) = w_0(x)$ (см. рисунок 1) и задача решена.

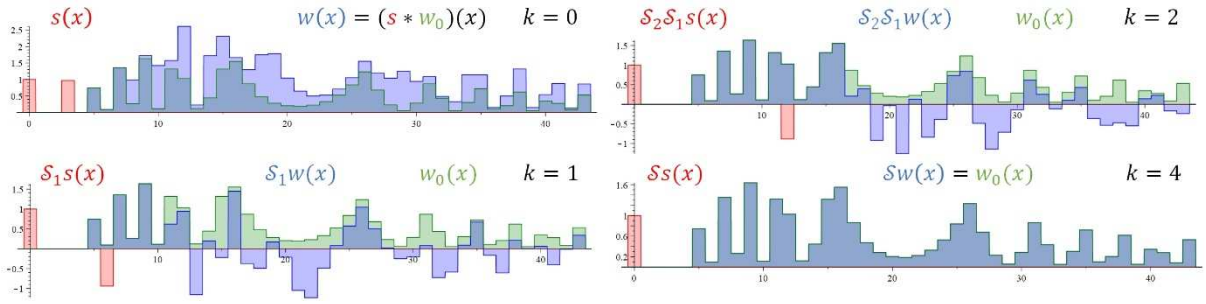


Рисунок 1 – Реконструкция сигналов при наложении двух импульсов

При наличии неопределенностей $\eta(x)$ погрешность $\mathcal{S}_k \cdots \mathcal{S}_1 \eta(x)$ на интервале $[2^{k-1}T, 2^k T)$ не превысит: $\eta(x)\left(1 + \sum_{i=1}^{k} a^{2^i}\right)$. При хорошей обусловленности задачи ($|a| < 1$) прирост погрешности a^{2^k} быстро уменьшается, а при плохой ($|a| > 1$) – возрастает (см. Таблицу 1).

Таким образом за $n = [log_2(L/T)] + 1$ итераций (или меньше, когда приростом погрешности можно пренебречь) можно на произвольном интервале $x \in (-\infty, L)$ реконструировать отклик любой ЛС-системы от одного импульса при наложении двух смещенных входных импульсов. При $v_0(x) = \delta(x)$ будет реконструирована передаточная функция $u(x)$ ЛС-системы. #

Таблица 1 – Прирост погрешности $\boldsymbol{a^{2^k}}$ на интервале $[\boldsymbol{2^{k-1}T}, \boldsymbol{2^kT})$ на $\boldsymbol{k}$-том шаге итерации

$k \setminus a$	**0.5**	**0.7**	**0.9**	**1.1**	**1.3**	**1.5**
1	0.25	0.49	0.81	1.21	1.69	2.25
3	$3.9\ 10^{-3}$	$5.8\ 10^{-2}$	$4.3\ 10^{-1}$	2.14	8.16	25.6
5	$2.3\ 10^{-10}$	$1.1\ 10^{-5}$	$3.4\ 10^{-2}$	21.1	$4.4\ 10^{3}$	$4.3\ 10^{5}$
7	$2.9\ 10^{-39}$	$1.5\ 10^{-20}$	$1.4\ 10^{-6}$	$1.9\ 10^{5}$	$3.8\ 10^{14}$	$3.5\ 10^{22}$
9	$7.5\ 10^{-155}$	$4.9\ 10^{-80}$	$3.7\ 10^{-24}$	$1.6\ 10^{21}$	$2.2\ 10^{58}$	$1.4\ 10^{90}$

3. Алгоритм численных вычислений

Алгоритм численных вычислений с помощью ПС-преобразований можно реализовать на основе решения задачи о произвольных дискретных наложениях импульсов.

Задача (дискретные наложения импульсов). Из отклика $w(x)$ ЛС-системы на входной сигнал $v(x) = (s * v_0)(x)$, где $s(x) = \sum_{i=0} s_i \delta(x - it)$, $s_0 \neq 0$, $t > 0$, $w(x) = v(x) = 0$ при $x < 0$, реконструировать на интервале $x \in (-\infty, L)$ отклик $w_0(x)$ от импульса $v_0(x)$.

Применяя к функции $s(x)$ ПС-преобразования $\mathcal{S} = \mathcal{S}_n \cdots \mathcal{S}_1$, где $\mathcal{S}_k$: $\delta(x) - (s_k^{k-1}/s_0)\delta(x - kt)$ и $\mathcal{S}_k \cdots \mathcal{S}_1 s(x) = s^k(x) = \sum_{i=0} s_i^k \delta(x - it)$, $s^0(x) = s(x)$, последовательно обнуляем коэффициенты s_k^k в функциях $s^k(x)$ (см. рисунок 2). На n-том шаге итерации при $nt > L$ на интервале $x \in (-\infty, L)$ получим: $s^n(x) = \mathcal{S}s(x) = s_0\delta(x)$. Отсюда, согласно (2): $\mathcal{S}w(x) = s_0 w_0(x)$. Погрешность решения на интервале $x \in [kt, kt + t)$ составит: $\mathcal{S}_k \cdots \mathcal{S}_1 \eta(x)/s_0$.

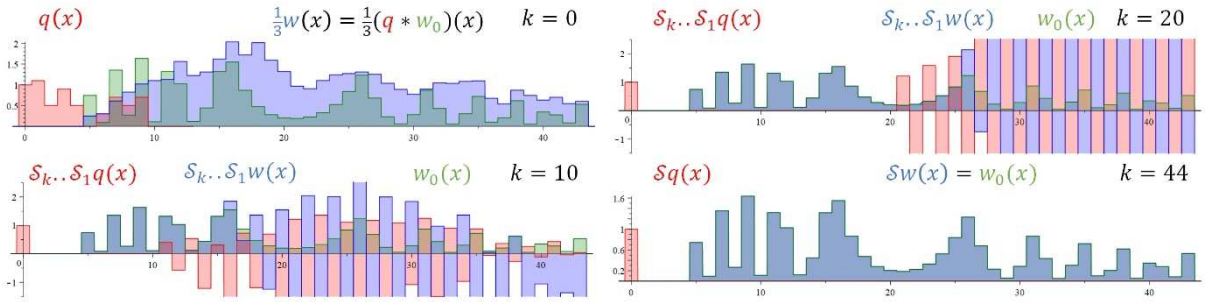


Рисунок 2 – Реконструкция сигналов при дискретных наложениях импульсов

Преимущество данного алгоритма заключается в том, что если на k-том шаге итерации $\left|s_k^{k-1}\right| < \varepsilon$, где ε – ошибка вычислений при численных расчетах или при наличии шумов и неопределенностей – ожидаемый уровень шумов в данных, то этот шаг можно пропустить и перейти к следующему, что значительно сокращает время вычислений, и снижает погрешность, т.к. неопределенности на пропущенных шагах итерации не добавляются.

В таблице 2 представлено время вычислений в секундах и отношения времен реконструкции сигналов длительности n при решении базовой задачи с помощью описанного алгоритма метода прямых сдвигов (ПС); обратного хода метода Гаусса-Жордана (ГЖ) $v_{i+1} = \left(w_i - \sum_{j=1}^{i} s_{i-j} w_j\right)/s_0$, $i = 0..n$; прямой матричной инверсии (ПМ) $v = \boldsymbol{S^{-1}}w$; а также LU- и SVD- методами, использующими соответствующие разложения матрицы $\boldsymbol{S}$. Расчеты проводились в среде Maple 2017.3 Waterloo Maple Inc. на ноутбуке HP 255 G7 с процессором AMD Ryzen 3, 2.5 ГГц. Использовались ПМ, LU и SVD методы из стандартных пакетов *LinearSolve* и *LinearAlgebra* программы Maple. Время вычислений с помощью метода прямых сдвигов значительно меньше времени вычислений другими методами. Например, при длине векторов $n = 6400$ время расчета с помощью ПС-метода в 2707 раз меньше времени расчета SVD-методом. Кроме того, при проведении расчетов ПС-методом никаких дополнительных массивов или матриц кроме исходных данных не требуется, что значительно экономит вычислительные ресурсы.

Таблица 2 – Время решения базовой задачи (сек) при различной длине сигналов $\boldsymbol{n}$

n	ПС	ГЖ	ПМ	LU	SVD
200	0.015	0.032 (2.13 ПС)	0.140 (9.33 ПС)	0.297 (19.8 ПС)	0.250 (16.7 ПС)
400	0.047	0.094 (2.00 ПС)	0.578 (12.3 ПС)	0.891 (19.0 ПС)	1.125 (23.9 ПС)
800	0.062	0.266 (4.29 ПС)	2.375 (38.3 ПС)	2.750 (44.4 ПС)	4.375 (70.6 ПС)
1600	0.094	1.250 (13.3 ПС)	11.38 (121.0 ПС)	11.70 (124.5 ПС)	22.47 (239.0 ПС)
3200	0.266	6.109 (23.0 ПС)	53.656 (201.7 ПС)	55.234 (207.6 ПС)	133.4 (501.5 ПС)
6400	0.719	23.97 (33.3 ПС)	294.06 (409.0 ПС)	329.44 (458.2 ПС)	1946.5 (2707.2 ПС)

Программные коды численных расчетов ПС-методом находятся в открытом доступе в библиотеке GitHub по адресу: https://github.com/novikov-borodin/lst-data-proc. При использовании программ ссылки на работы автора и директории автора в библиотеке GitHub обязательны.

4. Типовые задачи и примеры использования ПС-методов на практике

К задаче о наложении двух импульсов сводятся многие часто встречающиеся на практике задачи, которые будем называть типовыми. Задачи реконструкции отклика ЛС-системы от прямоугольного или трапецеидального импульса с равными фронтами из импульса большей длительности относятся к типовым, т.к. сдвигая сигналы и вычитая их из исходных, задача сводится к базовой. На рисунке 3 представлен пример реконструкции из отклика $w(t)$ емкостного делителя напряжения на трапецеидальный сигнал $v(t)$ длительности 120 нс его отклика $w_0(t)$ на трапецеидальный сигнал $v_0(t)$ длительности 30 нс. Представленные на графиках исходные сигналы были рассчитаны с помощью симулятора MicroCAP от Spectrum Software. На первом шаге ($k = 1$) задача сводится к базовой, и далее за две итерации на интервале от 0 до 500 нс сигналы были реконструированы с погрешностью 0.0012 %, которая была обусловлена в основном погрешностью дискретизации сигналов при численных расчетах, а не погрешностью ПС-метода.

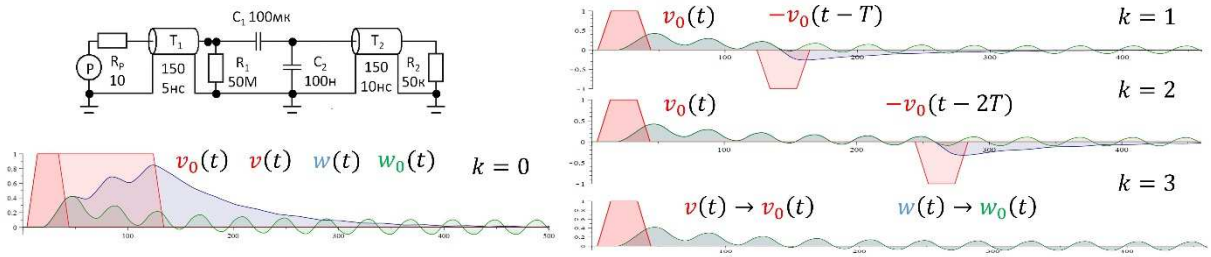


Рисунок 3 – Моделирование отклика делителя напряжения

Решение базовой задачи непосредственно использовалось в работе [4] для коррекции данных нейтронного спектрометра по времени пролета в Институте ядерных исследований РАН, где на основной импульс накладывался после-импульс, что приводило к искажению спектров. Эти искажения были успешно устранены всего за три итерации и были получены реальные спектры исследуемого вещества. Используя аналогичные рисунку 3 ПС-преобразования можно также смоделировать спектр подобных спектрометров от импульса меньшей длительности, что означает повышение их разрешающей способности без изменения конструкции.

На практике передаточные функции систем обработки часто описываются экспоненциальными функциями вида: $u(x) = Aexp(-x/B)$, $x > 0$. В дискретном виде с шагом разбиения Δ их можно представить как $u = A\sum_{i=0}\delta_i exp(-i\Delta/B)$, где δ_i – символ Кронекера. Всего за одну итерацию $\mathcal{S}_0$: $\delta_0 - \delta_1 exp(-\Delta/B)$ функцию u можно преобразовать в единичный импульс $\mathcal{S}_0 u = A\delta_0$ и реконструировать сигнал на входе из искаженного сигнала на выходе: $\mathcal{S}_0 w = (\mathcal{S}_0 u * v) = Av$ (см. рисунок 4а). В работах [5,6] передаточная функция ФЭУ представлялась в виде двух экспонент $u = A\sum_{i=0}\delta_i[exp(-i\Delta/B) - exp(-i\Delta/b)]$, $B > b > 0$, что позволило за два ПС-преобразования $\mathcal{S}_0$: $\delta_0 - exp(-\Delta/B)\delta_1$ и $\mathcal{S}_1$: $C[\delta_0 - exp(-\Delta/B)\delta_1]$, $C = [exp(-\Delta/B) - exp(-\Delta/b)]^{-1}$, реконструировать сигнал на входе из искаженного сигнала на выходе: $\mathcal{S}_1\mathcal{S}_0 w = (\mathcal{S}_1\mathcal{S}_0 u * v) = Av$ (см. рисунок 4б). Это решение используется как для регистрации, так и для разделения сигналов в детекторах частиц большой интенсивности.

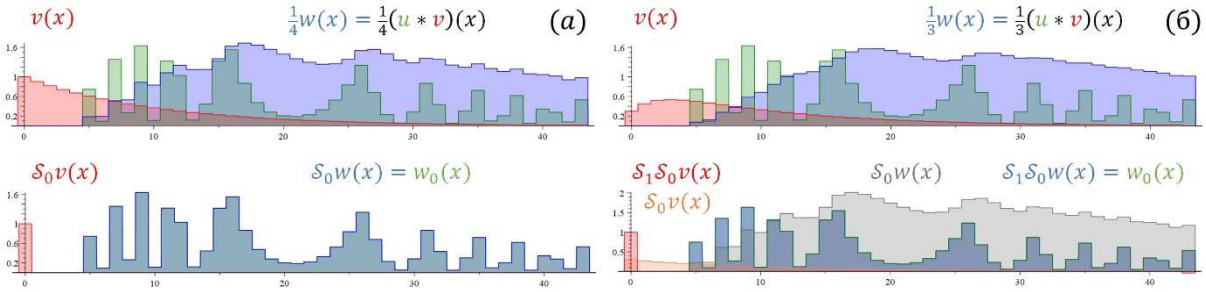


Рисунок 4 – Обработка экспоненциальных импульсов

Решения рассмотренных типовых задач для прямоугольных и экспоненциальных импульсов можно использовать для достижения сверх разрешения сигналов и изображений. Последовательно группируя n сигналов с разрешением T, смещенных относительно друг друга на шаг $t = T/n$, можно составить ЛС-уравнение (1), решением которого и будет сигнал с разрешением t. Причем для экспоненциальных импульсов сверх разрешение достигается всего за одну итерацию. В работе [7] ПС-преобразования использовались для моделирования заданных зон облучения $w(x)$ новообразований при протонной терапии. Под $v(x)$ понимался пик Брэгга, а реконструируемая функция $u(x)$ представляла собой искомый план облучения.

ПС-преобразования можно использовать для улучшения обусловленности задач. Типовые задачи плохо обусловлены, если при сведении их к базовой: $|s_0| < |s_1|$ (см. таблицу 1), что приводит к расходимости решений. Например, в случае трапецеидального сигнала с экспоненциальными фронтами на рисунке 5а, относительная погрешность реконструкции импульсного отклика $\|u_r - u\|/\|u\|$ с помощью ПС-алгоритма составляет 5688 %. Однако, если применить вспомогательное ПС-преобразование дифференцирования $\mathcal{S}: d/dx$ к уравнению $w(x) = (u * v)(x)$ и с помощью того же ПС-алгоритма решать задачу: $\mathcal{S}w(x) = (u * \mathcal{S}v)(x)$ (см. рисунок 5б), то из-за улучшения обусловленности задачи погрешность ее решения не превысит $1.52 \cdot 10^{-12}$ %.

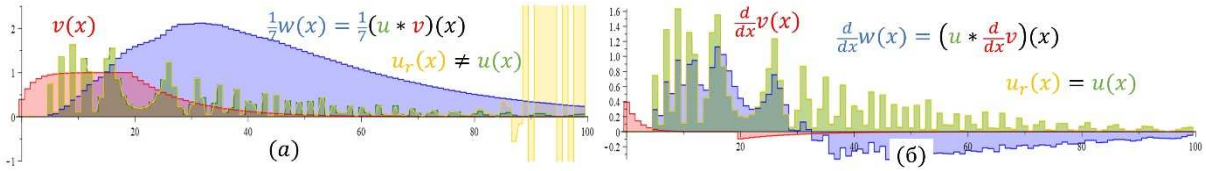


Рисунок 5 – Вспомогательные преобразования при решении плохо обусловленных задач

Рассмотренные типовые задачи и применения ПС-преобразований не являются единственными. Существует множество других преобразований и задач, решения которых можно свести к решению базовой задачи.

Заключение

Рассматриваемые в данной работе математические методы прямых сдвигов предназначены для коррекции, реконструкции и моделирования сигналов и изображений, искаженных в процессе обработки линейными стационарными системами, удовлетворяющими уравнению типа свертки. ПС-методы практически не ограничивают класс обрабатываемых функций и в целом ряде важных и часто встречающихся на практике случаях позволяют решать задачи обработки сигналов при минимуме исходной информации с минимальными погрешностями и вычислительными ресурсами, когда многие из существующих как прямых, так и приближенных математических методов оказываются менее эффективными. Эффективность ПС-методов обусловлена алгоритмом их реализации с возможностью выбора вида преобразований и сокращения их числа. ПС-методы имеют широкие перспективы практического применения и являются доступной альтернативой аппаратным методам, связанным с решением сложных научно-технических проблем и с созданием дорогостоящего оборудования.

Литература

1. Mueller J.L., Siltanen S. Linear and Nonlinear Inverse Problems with Practical Applications // SIAM, vol.10, 2012.
2. Hansen P.C. Discrete Inverse Problems: Insight and Algorithms. Fundamentals of Algorithms // Philadelphia: SIAM, 2010.
3. Hespanha J.P. Linear System Theory // Princeton university press, 2009.
4. Новиков-Бородин А.В. Реконструкция экспериментальных данных спектрометров по времени пролёта методом пошагового сдвига // Приборы и техника эксперимента, 2018, № 6, с. 17–24. https://doi.org/10.1134/S0032816218050269.
5. Jordanov V.T. Unfolding-synthesis technique for digital pulse processing, part 1: Unfolding // Nucl. Instrum. Methods Phys. Res. A 805, 63-71, 2016. doi: https://doi.org/10.1016/j.nima.2015.07.040
6. Jordanov V.T., Jordanova K.V. Unfolding-synthesis technique for digital pulse processing, Part 2: Synthesis // Nucl. Instrum. Methods Phys. Res. A 1004, 167421, 2022. https://doi.org/10.1016/j.nima.2022.167421.
7. Novikov-Borodin A.V. Direct Reconstruction of the Experimental Data in the Case of Ill-Conditioned Problems and in the Presence of Data Distortions // Instruments and Experimental Techniques, Vol.63, No.5, pp.658-669, © Pleiades Publishing, Ltd., 2020. doi: https://doi.org/10.1134/S0020441220050322.

APPLICATION OF DIRECT SHIFT METHODS FOR SOLVING THE PRACTICAL PROBLEMS OF ONE-DIMENSIONAL SIGNAL PROCESSING

A.V. Novikov-Borodin
Institute for Nuclear Research of RAS, 117312 Moscow, Russia

Abstract. The possibilities of using mathematical methods of direct shifts for correction, reconstruction and modeling of signals distorted by linear stationary systems during processing are considered. Linear stationary systems cover a wide class of processing systems described by convolution type equations. Direct shift methods make it possible, in a number of important and frequently encountered practical cases, to solve signal processing problems with minimal initial information with minimal computational resources and errors. The effectiveness of the direct shift method is due to the possibility of choosing the type of transformations and reducing their number. Examples of using the direct shift method in practice are considered. The results of comparing the direct shift method with Gauss-Jordan methods, direct matrix inversion, and LU and SVD decompositions are presented.

Program codes are openly available in GitHub library at: https://github.com/novikov-borodin/lst-data-proc.

Keywords: mathematical signal processing; direct shift techniques; linear stationary systems; convolution type equation.

ИССЛЕДОВАНИЕ ПОГРЕШНОСТИ ИЗМЕРЕНИЙ КОЭФФИЦИЕНТА ГАРМОНИК ГЭТ 188-2010 НА ЧАСТОТАХ 1—10 ГЦ

инж. Стальнова К.А., к.т.н. Безденежных С.В.

ФГУП «Всероссийский научно-исследовательский институт физико-технических и радиотехнических измерений» (ВНИИФТРИ)

Работа описывает исследования в рамках модернизации государственного первичного эталона ГЭТ 188-2010, направленной на расширение его частотного диапазона до 0.1 Гц. Для обеспечения точности измерений коэффициента гармоник (КГ) в новом диапазоне проведен анализ ключевых источников погрешности: неравномерности амплитудно-частотной характеристики и собственных нелинейных искажений измерительного тракта. Собственные искажения оценены двумя методами: прямыми измерениями с генератором и косвенным методом на основе анализа интермодуляционных искажений. Сравнение результатов, полученных в диапазоне 1–10 Гц, показало их общую согласованность, что подтверждает корректность методик. Выявленные расхождения на отдельных частотах требуют дополнительного изучения. Результаты работы закладывают основу для обеспечения требуемой точности эталона после его модернизации.

.В рамках модернизации государственного первичного эталона ГЭТ 188-2010 планируется расширение его частотного диапазона вниз – до 0,1 Гц вместо 10 Гц. Это позволит проводить высокоточные измерения коэффициента гармоник в сверхнизкочастотной области, что особенно важно для таких сфер, как: испытания материалов и вибродиагностика, геофизика и сейсмоакустика, гидроакустика. На промежуточном этапе модернизации (до частоты 1 Гц) выполняется комплекс работ по определению и учету собственных погрешностей эталона, что обеспечит необходимую точность измерений в новом диапазоне.

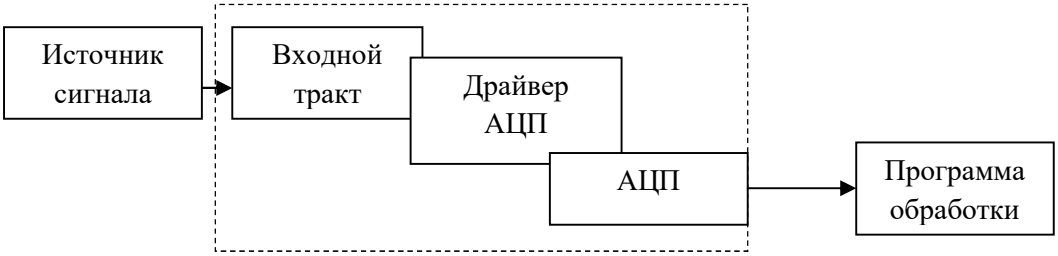


Рисунок 1. – Структурная схема измерений коэффициента гармоник ИКГЭ

Исследуемый сигнал оцифровывается, коэффициент гармоник определяется в результате обработки цифрового эквивалента исследуемого сигнала[1]. $K_г$ рассчитывается по следующей формуле:

$$K_г = 100 \times \frac{\sqrt{\sum_2^{10} U_i^2}}{U_1} \qquad (1),$$

где U_i – напряжения гармоник (учитывается 10 гармонических составляющих сигнала, включая основной тон).

Для обеспечения точности измерений амплитуды гармоник, полученные из спектра, корректируются с помощью коэффициентов, компенсирующих неравномерность амплитудно-частотной характеристики (АЧХ) измерительного тракта. Окончательное значение Кг вычисляется по формуле (1) с использованием этих скорректированных амплитуд.

Основными источниками суммарной погрешности измерения Кг в данном случае являются:

1. Погрешность из-за неравномерности амплитудно-частотной характеристики;
2. Погрешность из-за нелинейности коэффициента передачи от входа до АЦП(собственные искажения ИКГЭ);
3. Погрешность метода измерений Кг;

В рамках данной работы проведен анализ и количественная оценка двух ключевых источников погрешности: неравномерности АЧХ и собственных нелинейных искажений тракта.

Ошибка из-за неравномерности амплитудно-частотной характеристики

Для оценки погрешности, обусловленной неравномерностью амплитудно-частотной характеристики (АЧХ), измерения проводились на частотах первой гармоники 1, 3, 5, 7 Гц. Для каждой частоты первой гармоники АЧХ измерялась на 10 гармонических частотах. Измерение АЧХ измерителя Кг эталонного(ИКГЭ) производилось с помощью мультиметра 3458A и генератора DS360. Каждое значение, представленное в таблице 1, получено как среднее по десяти отсчётам при амплитуде входного сигнала 1 В. Программа управления ИКГЭ позволяет производить коррекцию АЧХ в соответствии с рассчитанными коэффициентами, при дальнейших измерениях и определении других составляющих погрешности это будет использоваться для исключения данной составляющей погрешности.

Таблица 1. Рассчитанные корректирующие коэффициенты АЧХ

Частота, Гц	Изм. Амплитуда, эталон, В	Амплитуда, 3458A, В	СКО напряжения, В	Взвешенные коэффициенты
1	1,14522	0,99915	0,000001	1
2	1,14506	0,99902	0,000008	1,00002
3	1,14510	0,99905	0,000005	1,00001
4	1,14514	0,99909	0,000005	1,00002
5	1,14539	0,99929	0,000006	1,00000
6	1,14580	0,99964	0,000003	0,99999
7	1,14624	1,00006	0,000004	1,00002
8	1,14676	1,00052	0,000006	1,00003
9	1,14752	1,00116	0,000006	1,00001
10	1,14845	1,00200	0,000005	1,00004

Собственные искажения

Для оценки собственных нелинейных искажений измерительного тракта были использованы два метода. Первый метод основан на прямых измерениях с использованием генератора сигналов со сверхнизкими искажениями на частотах 1—10 Гц. В качестве источника сигнала использовался генератор DS360, уровень нелинейных искажений которого согласно спецификации не превышает -106 дБ (Кг не более 0,0005 %)[2]. Учитывая, что собственные искажения генератора на порядок ниже ожидаемого уровня искажений исследуемого тракта, их вклад может быть признан пренебрежимо малым, следовательно, можно принять измеренные значения коэффициента гармоник за собственные. Результаты измерений в таблице 2.

Таблица 2. Измеренные значения Кг ИКГЭ

Частота основной гармоники, Гц	1	2	3	4	5	6	7	8	9	10
Измеренный КГ, %(СКО, %)	0,0059 (0,0009)	0,0027 (0,0001)	0,0027 (0,0002)	0,0034 (0,0002)	0,0037 (0,0002)	0,00279 (0,00011)	0,00239 (0,00013)	0,00214 (0,00012)	0,00236 (0,00012)	0,00244 (0,00016)

Второй способ заключается в косвенной оценке коэффициента гармоник (Кг) через анализ интермодуляционных искажений. Нелинейность коэффициента передачи тракта, которую в общем случае можно описать степенным рядом, приводит к появлению в выходном спектре не только гармоник исходных частот, но и комбинационных составляющих (интермодуляционных продуктов). Поскольку гармонические искажения (ответ на один тон) и интермодуляционные (ответ на два тона) порождаются одной и той же нелинейностью системы, между ними существует однозначная математическая связь [3]. Следовательно, выполнив спектральный анализ отклика на двухтональный сигнал и измерив амплитуды комбинационных составляющих, можно оценить коэффициент гармоник для случая одногармонического воздействия.

В общем случае для нелинейности n-го порядка комбинационные частоты и их амплитуды могут быть выражены через коэффициенты разложения в степенной ряд и амплитуды исходных гармоник. Для дальнейших расчётов примем, что с применением корректировочных

коэффициентов АЧХ на интересующем нас диапазоне частот равномерна. Для упрощения вычислений также ограничимся 5 гармониками (как показывают измерения, вклад гармоник свыше 5 ничтожно мал).

Так, для нелинейности 5-го порядка и входного сигнала $x(t) = \cos(\omega t)$ передаточную функцию можно выразить через степенной ряд, используя полиномы Чебышева первого рода. Полиномы Чебышева первого рода $T_n(x)$ связаны с гармониками косинуса следующим образом: $T_n(cos\theta) = \cos(n\theta)$. Для входного сигнала $x(t) = A\cdot\cos(\omega t)$ выходной сигнал $y(t)$ нелинейной системы 5-го порядка можно представить в виде:

$$y(t) = c_1\cdot\cos(\omega t) + c_2\cdot\cos(2\omega t) + c_3\cdot\cos(3\omega t) + c_4\cdot\cos(4\omega t) + c_5\cdot\cos(5\omega t); \quad (2)$$

используя для записи $x(t) = A\cdot\cos(\omega t)$ и упростив тригонометрические выражения получим:

$$y(t) = ax(t) + bx(t)^2 + cx(t)^3 + dx(t)^4 + ex(t)^5 \quad (3)$$

где коэффициенты a, b, c, d и e выражаются через c_n: $a = c_1 - 3c_3 + 5c_5$, $b = 2c_2 - 8c_4$, $c = 4c_3 - 20c_5$, , $d = 8c_4$, , $e = 16c_5$. Или $c_1 = a + 0{,}75\mathrm{c} + 0{,}625e, c_2 = 0{,}5b + 0{,}5d$, $c_3 = 0{,}25\mathrm{c} + 0{,}3125e, c_4 = 0{,}125d$ *и* $c_5 = 0{,}0625e$ (амплитуды 1—5 гармоник соответственно).

Рассмотрим в качестве входного сигнала сумму 2 гармонических сигналов. Т.е. перейдем к анализу интермодуляции, возникающей при той же самой нелинейности. Выходной сигнал y(t) будет содержать следующие частотные компоненты—основные частоты: ω_1, ω_1,; гармоники($2\omega_1$,$2\omega_2$,…, $5\omega_1$,$5\omega_2$), интермодуляционные продукты ($\omega_1 \pm \omega_2$,…,$3\omega_1 \pm 2\omega_2$, $3\omega_2 \pm 2\omega_1$,

Нелинейность 5-го порядка описывается степенным рядом:

$$y(t) = \sum_{n=0}^{5} c_n\, x(t)^n \quad (5)$$

где c_n— коэффициенты, определяющие вклад каждой степени $x(t)$. Подставим $x(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$ и раскроем степени(ниже представлены выражения для некоторых членов):

Линейный член (n=1): $x(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$

Пятая степень (n=5):

$$x(t)^5 = \big(cos(\omega_1 t) + cos(\omega_2 t)\big)^5$$

$$\begin{gathered} x(t)^5 = 6{,}25cos(\omega_1 t) + 6{,}25cos(\omega_2 t) + 1{,}5625cos(3\,\omega_1 t)\,) + 1{,}5625cos(3\,\omega_1 t) + \\ +3{,}125cos((2\omega_1 - \omega_2)t + 3{,}125cos((2\omega_2 - \omega_1)t + 3{,}125cos((2\omega_1 + \omega_2)t + \\ +3{,}125cos((2\omega_2 + \omega_1)t + 0{,}625cos((3\omega_2 - 2\,\omega_1)t + 0{,}625cos((3\omega_2 + 2\,\omega_1)t + \\ +0{,}3125cos(4\omega_2 t + \omega_1) + 0{,}3125cos(4\omega_2 t - \omega_1) + 0{,}3125cos(4\omega_1 t + \omega_2) + \\ + 0{,}3125cos(4\omega_1 t - \omega_2) + 0{,}625\,cos((3\omega_1 + 2\,\omega_2)t + +0{,}625\,cos\big((3\omega_1 - 2\,\omega_2)t\,\big) + \\ +0{,}0625\,cos(5\omega_1 t) + 0{,}0625\,cos(5\omega_2 t) \end{gathered} \quad (6)$$

Результирующее выражение с учетом коэффициентов a,b,c,d, e и группировкой:

$$\begin{gathered} y(t) = (2{,}25c + 6{,}25e)\cdot cos(\omega_1 t) + (2{,}25c + 6{,}25e)\cdot cos(\omega_2 t) + (0{,}5b + 2d)\cdot cos(2\omega_1 t) + \\ + (0{,}5b + 2d)\cdot cos(2\omega_2 t) + (b + 3d)\cdot cos((\omega_1 + \omega_2)t) + (b + 3d)\cdot cos\big((\omega_1 - \omega_2)t\big) + \\ + (0{,}25c + 1{,}5625e)\cdot cos(3\omega_1 t) + (0{,}25c + 1{,}5625e)\cdot cos(3\omega_2 t) + \\ +(\,0{,}75c + 3{,}125e)\cdot cos\big((2\omega_1 + \omega_2)t\big) + (\,0{,}75c + 3{,}125e)\cdot cos\big((2\omega_1 - \omega_2)t\big) + \\ +(\,0{,}75c + 3{,}125e)\cdot cos\big((2\omega_2 + \omega_1)t\big) + (\,0{,}75c + 3{,}125e)\cdot cos\big((2\omega_2 - \omega_1)t\big) + \\ +0{,}5d\cdot cos\big((3\omega_1 - \omega_2)t\big) + 0{,}5d\cdot cos\big((3\omega_2 - \omega_1)t\big) + 0{,}5d\cdot cos\big((3\omega_1 + \omega_2)t\big) + \\ + 0{,}5d\cdot cos\big((\omega_1 + 3\omega_2)t\big) + 0{,}125d\cdot cos(4\omega_2 t) + 0{,}125d\cdot cos(4\omega_2 t) + \\ + 0{,}75d\cdot cos\big((2\omega_1 - 2\omega_2)t\big) + 0{,}75d\cdot cos\big((2\omega_1 + 2\omega_2)t\big) + \\ + 0{,}625e\cdot cos\big((3\omega_1 - 2\omega_2)t\big) + 0{,}625e\cdot cos\big((3\omega_1 + 2\omega_2)t\big) + \\ + 0{,}3125e\cdot cos\big((4\omega_2 + \omega_1)t\big) + 0{,}3125e\cdot cos\big((4\omega_1 + \omega_2)t\big) + \\ + 0{,}625e\cdot cos((5\omega_2)t) + 0{,}625e\cdot cos((5\omega_1 t) \end{gathered} \quad (7)$$

Далее при помощи (3) можно вычислить, как соотносятся коэффициент гармоник и амплитуды соответствующих комбинационных частот.

Для экспериментальной проверки косвенного метода был реализован двухтональный тест. Его суть заключается в подаче на вход измерительного тракта двух синусоидальных сигналов с

близкими частотами, с двух генераторов через смеситель. Для соответствия однотональному воздействию с амплитудой A амплитуда каждого из двух тонов устанавливается равной $A/\sqrt{2}$, что обеспечивает равенство средней мощности сигналов. Измеряемый сигнал оцифровывается 24-разрядным АЦП AD7760, исходная частота дискретизации составляет 20 МГц, За счёт передискретизации в 256 раз эффективная частота дискретизации равна 78125 Гц. Максимальная длина выборки составляет 1048576 24-разрядных отсчета. Для обработки экспериментальных данных разработана программа на языке Python с графическим интерфейсом (библиотека Tkinter). Программа загружает файл с отсчётами сигнала и с помощью алгоритма Герцеля вычисляет амплитуды на заданных комбинационных частотах. Эти значения используются затем для расчета коэффициента гармоник согласно математической модели. Применение программы автоматизирует расчёты, повышает достоверность результатов и исключает ручные ошибки при обработке.

Таблица 3. Результаты вычислений значения Кг ИКГЭ

Частота основной гармоники, Гц	1	2	3	4	5	6	7	8	9	10
Полученный КГ, %(эталон, ПО)	0,0035	0,0026	0,0027	0,0027	0,0024	0,0028	0,0021	0,0019	0,0019	0,0017
Полученный КГ, %(DS360)	0,0059	0,0027	0,0027	0,0034	0,0037	0,0028	0,0024	0,0021	0,0024	0,0024

Результаты исследований показывают согласованность данных, значения коэффициентов гармоник, полученные двумя методами, находятся в пределах одного порядка величин. Полученные данные подтверждают корректность методик в целом, однако выявленные расхождения на частотах 1 Гц и 5 Гц требуют дальнейшего уточнения и анализа причин, среди которых наиболее вероятными являются влияние фона частоты сети и особенности аналогового тракта.

Литература

1. Платонов Ф. А., Ахромеева Е. О., Бабарыкин В. А., Пругло А. В., Безденежных С. В., Каминский О. В., Стальнова К. А. Государственный первичный эталон единицы коэффициента гармоник в диапазоне (0,001...100) % для сигналов с основной гармоникой в диапазоне частот (10...200000) Гц ГЭТ 188-2010 // Измерительная техника. 2016. № 3. С. 3–6.
2. Low Distortion Function Generator DS360 [Электронный ресурс] // Stanford Research Systems. – URL: https://www.thinksrs.com/products/ds360.html.
3. Богданович, Б. М. Нелинейные искажения в приемно-усилительных устройствах — М.: Связь, 1980. —280 с.

RESEARCH OF THE ERROR OF MEASURING HARMONIC COEFFICIENT VALUES OF NATIONAL STANDARD OF THE HARMONIC COEFFICIENT UNIT AT FREQUENCIES OF 1 - 10 HZ

eng. Stalnova K.A., Ph.D. Bezdeneznykh S.V.

Russian Metrological Institute of Technical Physics and Radio Engineering (VNIIFTRI)

The article presents research conducted as part of the modernization of the State primary standard of the harmonic coefficient unit of the Russian Federation, aimed at expanding its frequency range to 0.1 Hz. To ensure accurate measurements in this new range, an analysis of key sources of error was conducted: unevenness of the amplitude-frequency response and the inherent nonlinear distortions of the measuring path. Intrinsic distortions were estimated using two methods: direct measurements with a generator and an indirect method based on intermodulation distortion analysis. A comparison of the results obtained in the 1–10 Hz range showed their overall consistency which confirms the validity of the methods. The identified discrepancies at individual frequencies require further study.

ПРИМЕНЕНИЕ МЕТОДА ЛИНЕЙНОЙ РЕГРЕССИ ДЛЯ ОПРЕДЕЛЕНИЯ ХАРАКТЕРИСТИК ПОЛИГАРМОНИЧЕСКИХ СИГНАЛОВ

ктн. Шерстобитов С.В.[1]

[1] ФГУП «ВНИИФТРИ»

В статье описано применение методов линейной регрессии при цифровой обработке выходных сигналов АЦП для определения характеристик полигармонических сигналов. Приведены результаты статистического моделирования, которое использовалось для определения точностных характеристик метода. Моделирование проводилось на цифровых реализациях сигналов при различных значениях разрядности АЦП, отношения сигнал/шум и от числа обрабатываемых отсчетов. Показано, что при применении данного метода обработки сигналов цифровых реализаций возможно решение задач высокоточного определения характеристик измеряемых сигналов переменного напряжения и полигармонических сигналов, а также определение динамических характеристик АЦП.

Постановка задачи

За последние 20 лет значительно расширилась область применения измерительных приборов, использующих в своем составе высокоточные АЦП и ЦАП. Сами эти преобразователи значительно развились в части быстродействия, точности, разрешения и энергопотребления. Это приводит к необходимости улучшить реализацию существующих эталонов, чтобы было возможным проводить калибровку устройств на их основе на производстве и в процессе эксплуатации для достижения максимально возможной точности. В передовых мировых технологиях используются программируемые установки напряжения на Эффекте Джозефсона PJVS (Programmable Josephson Voltage Standard) [1], установки на основе импульсной накачки JAWS (Josephson Arbitrary Waveform Synthesizer) и другие, эти установки так же по своему принципу действия являются цифро-аналоговыми преобразователями, только работающими с практически «фундаментальной точностью» и «абсолютной линейностью» [2]. Методика определения характеристик ЦАП-АЦП в целом подходит и для определения метрологических характеристик установок на эффекте Джозефсона.

Технологический уровень микроэлектроники в настоящее время позволил создать АЦП с очень высокими характеристиками: разрядностью 16 бит, 24 бит и более, вплоть до 32 бит с частотой дискретизации, достигающей сотен кГц и даже сотен МГц (при меньшей разрядности). Естественно, что процесс развития АЦП и улучшения их характеристик сопровождается соответствующим изменением функций, выполняемых этим блоком. Изменяется также постановка задачи определения МХ АЦП. Так для АЦП, работающих в режиме преобразования постоянных уровней сигнала, основными определяемыми характеристиками являются точности преобразования этих уровней, т.е. статические характеристики. Следующим этапом является работа АЦП на входных сигналах, медленно меняющихся во времени. Соответствующим шагом стало определение МХ АЦП на сигналах, скачкообразно изменяющихся во времени. В этом случае возникла задача измерения динамических характеристик АЦП, таких как различные времена задержки, время установления уровня выходного сигнала и т.п. Эти характеристики по сути своей являются релаксационными характеристиками блока АЦП.

В таком случае необходим переход к рассмотрению АЦП, как стационарной динамической системы с определёнными динамическими характеристиками. Набор динамических характеристик, определяющих качество работы АЦП при входном сигнале, изменяющемся во времени, зависит от назначения АЦП, т.е. от класса входных сигналов как функций времени и от дальнейшего использования выходного цифрового сигнала. Необходимо отметить, что этот класс должен входить в класс рабочих входных сигналов блока АЦП. Для АЦП, осуществляющих преобразование отдельных установившихся значений сигнала, классом входных сигналов является класс функций, принимающих постоянные значения, и характеристиками АЦП в таком режиме будут статические характеристики. Для работы различных АЦП в динамическом режиме характерно разнообразие классов входных сигналов и описывающих их систем параметров. Поэтому невозможно ввести общие динамические характеристики качества АЦП для всех возможных классов входных сигналов и режимов работы.

Практически процесс определения значений характеристик АЦП может быть реализован в следующей последовательности:

1) Подача на вход АЦП эталонного аналогового тестового сигнала, входящего в класс допустимых сигналов, с заданными значениями параметров;

2) Регистрация полученной в рабочем режиме реального времени последовательности цифровых отсчётов в запоминающее устройство;

3) Вычисление значений параметров сигнала путём компьютерной обработки записанной последовательности цифровых отсчётов с применением специального программного обеспечения;

4) Сравнение полученных значений параметров с заданными и вычисление значений характеристик АЦП как погрешностей преобразования сигнала.

Для реализации предлагаемого принципа определения характеристик АЦП оптимальным выбором тестового сигнала является синусоидальный сигнал. Ниже приведено обоснование такого выбора:

1) Синусоидальный сигнал входит в общий класс сигналов, имеющих спектр в определённой полосе частот. Этот класс достаточно широк: в него входят такие сигналы, как узкополосные, импульсные, видеосигналы и др.

2) Специальные виды сигналов, в которых полезная информация "зашита" каким-либо сложным путём (использование особым способом модулированных несущих и поднесущих, импульсно-кодовая модуляция и др.) при прохождении через определённые цепи и преобразователи во многих случаях рассматриваются как сигналы, состоящие из синусоидальных составляющих с частотами в определённой полосе.

3) Для генерации аналогового синусоидального сигнала существуют генераторы, имеющие достаточно высокие характеристики точности, превосходящие характеристики генераторов других типов сигналов.

У разработанного в лаборатории ФГУП ВНИИФТРИ ПО, работа алгоритма программы основана на методе регрессионного анализа по принципу наименьших квадратов.

Целью проводимых исследований являлось создание и экспериментальная апробация ПО для выполнения вышеизложенных функций в классе гармонических сигналов с погрешностью измерения переменного напряжения, по крайней мере, не превышающей в относительных единицах $10^{-6}…10^{-7}$.

Описание алгоритмов работы программного обеспечения

При цифровой обработке сигналов часто используются алгоритмы БПФ, ДПФ, которые имеют свои ограничения по применению и ограничивают погрешность измерения. Построение высокоточных генераторов тестовых сигналов на базе Джозефоновского синтезатора сигналов произвольной формы и разработка методик определения характеристик АЦП и характеристик полигармонических сигналов при помощи АЦП является актуальной современной задачей [3]. В лаборатории ФГУП «ВНИИФТРИ» были разработаны алгоритмы для определения характеристик гармонических сигналов и характеристик ЦАП/АЦП с использованием метода регрессионного анализа. Разработано программное обеспечение, которое может быть использовано для генерации цифровых реализаций тестовых сигналов и определения характеристик сигналов переменного напряжения или характеристик АЦП.

Для решения этой задачи были проведены следующие работы:

1. Был разработан специальный программный модуль - генератор цифрового тестового сигнала (ГЦТС). Такой генератор вычисляет массив последовательных цифровых отсчётов сигнала АЦП в соответствии с заданной моделью тестового сигнала из класса рабочих сигналов. В полной модели выходного сигнала АЦП учитываются искажения сигнала при его преобразовании АЦП (нелинейность, аддитивный шум и т.д.). Задаются значения всех параметров исходной модели сигнала и параметров АЦП. Вычисленные модельные значения отсчётов преобразуются в формат выходного сигнала АЦП и записываются в файл данных. Создаётся также файл с записью значений всех задаваемых параметров модели сигнала.

Модуль генератора формирует цифровой сигнал, соответствующий гармоническому сигналу

$$A(t,N) = A_0 + \sum_{l=1}^{M} A_l \sin(2\pi \mathrm{l} \mathrm{F}_0 + \Psi_l) + B(t,\sigma) \qquad (1)$$

где A_0 - постоянная составляющая сигнала $A\ (t,\ N)$; A_l, $2\pi lF_0$ и $\underline{\psi_l}$ - амплитуда, угловая частота и фаза l-й гармоники составляющей $A(t,\ N)$ синусоидального сигнала; M - количество гармоник основного тона синусоидальной составляющей в сигнале $A(t,\ N)$, $B(t,\sigma)$ является случайной функцией, описывающей аддитивный "белый" шум с нулевым средним значением и стандартным отклонением, равным σ.

2. Разработан дополнительный программный модуль для обработки выходного цифрового сигнала АЦП, осуществляющий после расчета с помощью исследуемого ПО вычисленных значений параметров сигнала, включая параметры искажений, путём сравнения полученных и заданных значений параметров и определяющий точностные характеристики работы ПО.

Для обработки цифровых реализаций используется метод регрессионного анализа. Внутренним масштабным параметром реализации является количество отсчетов за период основного тона. При известном значении задача определения параметров синусоидального сигнала формулируется как линейная регрессия по принципу наименьших квадратов в классе сумм тригонометрических функций. Затем алгоритм вычисляет амплитуды и фазы основного тона и гармоник, коэффициент нелинейных искажений, эффективное число разрядов АЦП, отношение сигнал/шум, динамический диапазон и т.д.

Результаты исследований характеристик программного обеспечения

Для определения собственных погрешностей, вносимых ПО при определении частотных и амплитудных параметров, был использован процесс статистического моделирования, состоящий из циклов моделирования и обработки реализаций, статистической обработки.

Исследование характеристик ПО проводилось путём их непосредственного определения по модельным реализациям сигналов с известными значениями задаваемых параметров.

Для исследования характеристик разработанного ПО использовались тестовые сигналы, содержащие основной тон (основная частота) и аддитивный белый шум. Число обрабатываемых реализаций для каждого значения параметров составило от 500 до 1000, что обеспечивает не менее двух верных десятичных знаков в полученных оценках. Результаты исследований и значения параметров, при которых они проводились, представлены ниже в виде таблиц 1 и 2 и графиков на рисунках 1 и 2. Исследовались предельные точностные характеристики ПО, которые можно ожидать, в случае работы с ЦАП/АЦП или Джозефсоновскими преобразователями 9 Бит (динамический диапазон порядка 60 дБ) и 16 бит (динамический диапазон порядка 100 дБ).

Исследования проводились при следующих заданных значениях параметров сигнала:

K = 132,547 – число отсчётов на период основного тона сигнала.

N = 500-100000 – число отсчётов в реализациях.

Таблица 1. Результаты исследования программного обеспечения при использовании 9-битного АЦП. (Динамический диапазон 60 дБ).

	N (число обрабатываемых отсчётов)					
	500	1000	3000	10000	30000	100000
Amp1	9.12e-05	6.36e-05	3.78e-05	1.94e-05	1.19e-05	5.64e-06
Phas1	5.45e-05	4.16e-05	2.3e-05	1.3e-05	9.26e-06	3.98e-06
Gamma	1.3e-05	4.6e-06	8.84e-07	1.53e-07	3.24e-08	4.08e-09

Таблица 2. Результаты исследования программного обеспечения при использовании 16-битного АЦП. (Динамический диапазон 96 дБ).

	N (число обрабатываемых отсчётов)					
	500	1000	3000	10000	30000	100000
Amp1	5.6e-07	3.95e-07	3.06e-07	1.28e-07	7.37e-08	3.77e-08
Phas1	3.52e-07	2.55e-07	1.93e-07	7.94e-08	4.54e-08	2.49e-08
Gamma	8.32e-08	2.88e-08	1.8e-08	9e-10	1.77e-10	2.9e-11

На рисунке 3 показаны графики оценки погрешности ПО при вычислении амплитуды, фазы и частоты сигналов в зависимости от отношения сигнал/шум от 60 до 140 дБ.

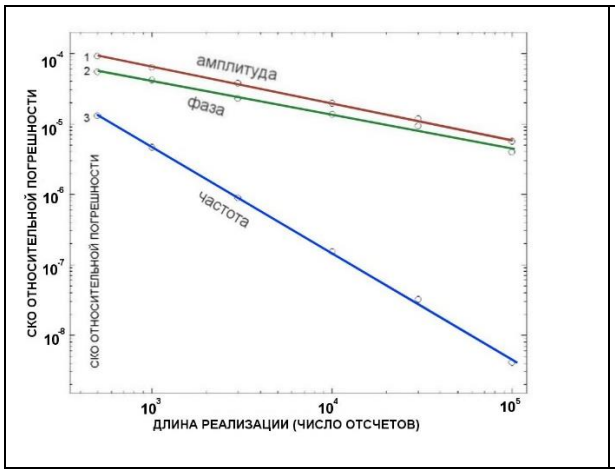


Рис. 1. Результаты исследования работы ПО при обработке сигналов АЦП 9 бит.
1 – Amp1 – (СКО) относительной погрешности вычисления амплитуды основного тона, красного цвета; 2 – Phas1 – (СКО) погрешности вычисления фазы основного тона (в единицах PI Period Interval) зеленый цвет; 3 – Gamma – (СКО) относительной погрешности вычисления частоты основного тона, синий цвет.

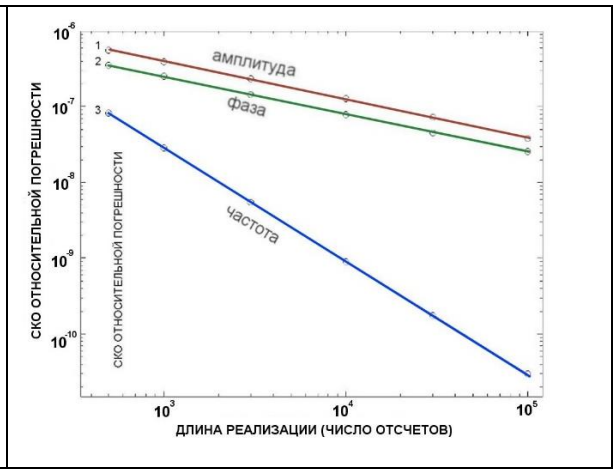


Рис. 2. Результаты исследования работы ПО при обработке сигналов АЦП 16 бит.
1 – Amp1 – (СКО) относительной погрешности вычисления амплитуды основного тона, красного цвета; 2 – Phas1 – (СКО) погрешности вычисления фазы основного тона (в единицах PI Period Interval) зеленый цвет; 3 – Gamma – (СКО) относительной погрешности вычисления частоты основного тона, синий цвет

На рисунке 4 показаны графики зависимостей относительной погрешности ПО, полученных на модельных реализациях при определении амплитуды и частоты сигнала от отношения сигнал/шум при числе обрабатываемых отсчетов 1000 и 10000. Из приведенных в логарифмическом масштабе графиков видно, что полученные зависимости таких характеристик, как СКО погрешностей основных параметров (амплитуды основного тона, фазы и частоты) от отношения сигнал/шум и от длины реализаций имеют линейный характер. Это соответствует аналитической зависимости между указанными параметрами в виде степенной функции. То есть для возможности более точного вычисления требуется повышать сигнал-шум, и обрабатывать большее число отсчетов сигнала.

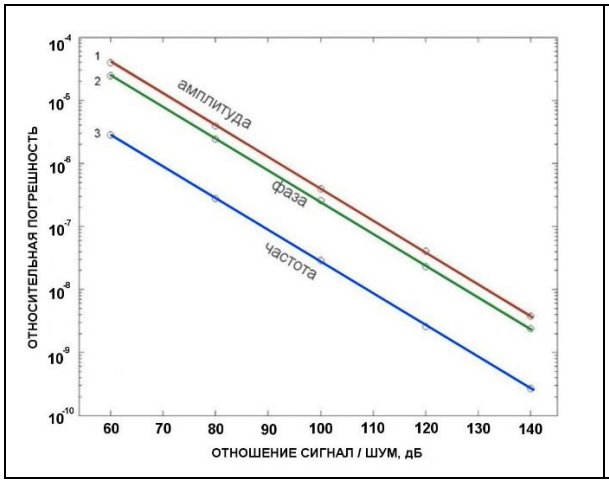


Рис. 3. Графики оценки погрешности ПО при вычислении амплитуды, фазы и частоты сигналов в зависимости от отношения сигнал/шум (число обрабатываемых отсчетов сигнала 1000) число отсчетов на период сигнала 132,547698

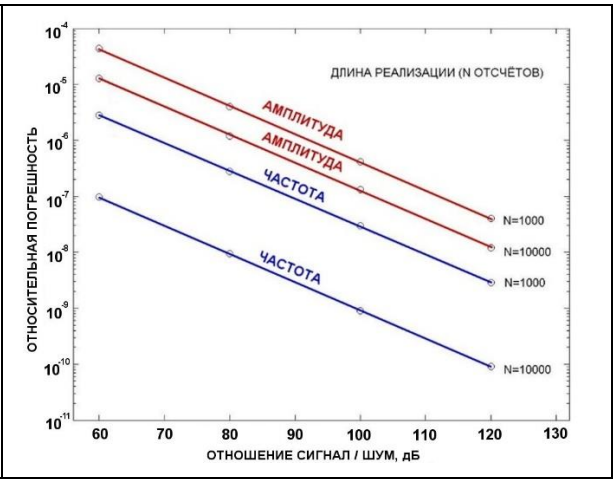


Рис. 4. Графики зависимости относительной погрешности ПО, полученных на модельных реализациях, при определении амплитуды (красные линии) и частоты сигнала (синие линии) от отношения сигнал/шум при числе обрабатываемых отсчетов 1000 и 10000.

Из графиков следует, что оценивая погрешность работы данного ПО для работы с 9 битными АЦП или более высоко разрядными преобразователями, но в условиях отношения сигнал/шум не более 60 дБ, при числе обрабатываемых отсчетов 1000, что соответствует приблизительно восьми периодам исследуемого сигнала, относительная погрешность определения амплитуды и фазы не превышает $7 \cdot 10^{-5}$ и $5 \cdot 10^{-5}$ соответственно, и частоты $5 \cdot 10^{-6}$ и эти погрешности снижаются при увеличении числа обрабатываемых отсчетов. Такой погрешности работы ПО может быть вполне достаточно для многих применений, не требующих сверх высокой точности или, например, в высокочастотной области, в которой эффективное число разрядов АЦП как правило не превышает 10-14 бит. Что может быть внедрено, например, для расширения функциональных технических возможностей современных цифровых осциллографов для анализа гармонических сигналов.

При использовании разработанного ПО для работы с 16 битными АЦП или более высоко разрядными преобразователями, но в условиях отношения сигнал/шум не превышающего 100 дБ, при тех же условиях, относительная погрешность разработанного ПО при определении амплитуды и фазы не превышает $4 \cdot 10^{-7}$ и $3 \cdot 10^{-7}$, соответственно, и частоты $3 \cdot 10^{-8}$, и также снижается при увеличении числа обрабатываемых отсчетов. Это соответствует уровню достижимой погрешности, к которому стремятся метрологи в случае проведения сличений эталонов переменного напряжения на эффекте Джозефсона в высшем звене метрологии электрических измерений.

Из полученных результатов проведенных исследований также следует, что разработанное ПО хорошо работает в широком диапазоне для числа отсчётов на период сигнала. Алгоритмы, заложенные в работу данного ПО, при обработке сигналов многоразрядных АЦП, жестко привязанных к эталону частоты и времени, могут быть использованы для высокоточного измерения не только амплитуды и фазы, но и частоты сигнала при относительно коротких выборках сигнала, либо на длинных выборках, с уменьшенной погрешностью, вносимой алгоритмом ПО.

При использовании в аналогичных целях более простых методов преобразования Фурье и подхода Блекмана-Тюльки, можно сказать, что они не дают необходимой точности, особенно при обработке зашумленных сигналов. Например, методам на основе преобразования Фурье свойственны следующие недостатки: из-за эффекта конечной длины обрабатываемых данных происходит «размывание» частотной информации что приводит к увеличению погрешности измерения частоты, ограничение частотного разрешения и связанного с ним отсутствия различения спектральных линий, составляющих сигнала, явления просачивания на соседние гармоники, что искажает спектр сигнала.

Для сравнения полученных результатов оценки собственной погрешности программного обеспечения с другими методами, можно привести для примера, алгоритм, описанный в публикации [4]. В статье [4] описан алгоритм определения таких же параметров сигналов как амплитуда и частота, основанный на методе разложения сигнала на собственные числа, который основан на анализе собственных значений автокорреляционной матрицы или одной из матриц данных, и результаты моделирования его работы для оценки метрологических характеристик. Этот метод по мнению автора публикации обеспечивает лучшие характеристики разрешения и оценивания параметров, чем другие параметрические методы, особенно при низких отношениях сигнал/шум, когда эти методы не в состоянии различить близкие по частоте синусоиды или другие узкополосные спектральные компоненты. В статье автор подтверждает, что для данного метода на основе разложения сигнала на собственные числа минимальная погрешность по амплитуде составляет 0,02 % ($2 \cdot 10^{-4}$ отн. ед.) и частоты 0,01 % ($1 \cdot 10^{-4}$ отн. ед.), которая достигается при 8-12 разрядов АЦП и при увеличении разрядности АЦП практически не изменяется.

Также в публикации делается вывод, что для стабильного определения параметров сигнала необходимо, чтобы количество периодов сигнала было больше пятидесяти, тогда погрешности определения параметров находятся на уровне 0,05% ($5 \cdot 10^{-4}$ отн. ед.) и 0,01% ($1 \cdot 10^{-4}$ отн. ед) для амплитуды и частоты соответственно. Если сравнить метод, основанный на разложении сигнала на собственные числа, с описанным выше и исследованным в данной работе методом на основе линейной регрессии, то второй при работе с 9 разрядными АЦП (отношение сигнал/шум 60 дБ) и числе обрабатываемых отсчетов 1000 дает предельные точностные характеристики

практически на порядок или минимум в 5 раз лучше при измерении амплитуды и в 20 раз при измерении частоты чем метод на основе разложения автокорреляционной матрицы сигнала на собственные числа. При увеличении разрядности АЦП до 16 и увеличении отношения сигнал/шум до 100 дБ у метода на основе линейной регрессии погрешность определения этих параметров уменьшается ещё практически на два порядка в то время, как для метода на основе разложения автокорреляционной матрицы сигнала на собственные числа погрешность определения амплитуды и частоты практически не изменяется. Нужно отметить, что автор метода на основе разложения автокорреляционной матрицы сигнала на собственные числа не позиционирует его как метод для сличения эталонов переменного напряжения, а скорее, как метод для использования в виртуальных измерительных приборах на основе персонального компьютера, оснащенного стандартным АЦП в качестве универсальной измерительной станции, или в микропроцессорах цифровой обработки сигналов, особенно в присутствии шумов.

Заключение

Результаты статистического моделирования с применением описанного алгоритма, опирающегося на методы линейной регрессии, при обработке выходных сигналов АЦП, показали возможные точностные характеристики, достигаемые при измерении таких параметров сигнала как: амплитуда, фаза, частота основного тона и гармоник, которые зависят от отношения сигнал/шум, длины обрабатываемой реализации. Погрешности метода при обработке реальных сигналов будут зависеть от точности применяемых АЦП, точности и стабильности сигнала частоты дискретизации.

При апробировании данного алгоритма при проведении измерений с установкой на эффекте Джозефсона, были получены подтверждающие расчётные значения результаты, возможно не достижимые при использовании других методов обработки, таких как БПФ, ДПФ, метод разложения сигнала на собственные числа. Планируется в дальнейшем применять эти алгоритмы при разработке методик измерения характеристик АЦП и характеристик синтезаторов переменного напряжения на эффекте Джозефсона для получения относительных погрешностей от единиц до долей на 10^{-6}.

Литература

1. C.A.Hamilton, C.J.Burroughs, R.L.Kautz, IEEE Trans. Instrum. Meas., 1995, 44, pp.223-225.
2. С.В. Шерстобитов «Квантовые цифровые синтезаторы постоянного и переменного электрического напряжения, практика использования», Альманах современной метрологии, № 2, 2015, с. 43-56.
3. M. Sira, O. Kieler, R. Behr, A Novel Method for Calibration of ADC Using JAWS, EEE Trans. Instrum. Meas, 2019, vol. 68, no. 6, p. 2091-2099; doi:10.1109/TIM.2018.2888918.
4. Козлов В.В. Определение параметров гармонических сигналов в условиях действия шумов и помех на основе метода разложения сигнала на собственные числа // Современные проблемы науки и образования. – 2013. – № 6.

APPLICATION OF LINEAR REGRESSION METHOD TO DETERMINE THE CHARACTERISTICS OF POLYHARMONIC SIGNALS

Sherstobitov S.V.[1]

[1] FSUE VNIIFTRI

The article describes the application of linear regression methods in the digital processing of ADC output signals to determine the characteristics of polyharmonic signals. The results of statistical modeling, which was used to determine the accuracy characteristics of the method, are presented. The modeling was conducted on digital implementations of signals with different ADC bit depths, signal-to-noise ratios, and the number of samples processed. It is shown that the application of this method for processing digital implementations of signals allows for the accurate determination of the characteristics of measured AC voltage and polyharmonic signals, as well as the determination of the dynamic characteristics of the ADC.

28-я Международная конференция «Цифровая обработка сигналов и её применение – DSPA-2026», Москва, Россия, доклады.

Серия: Цифровая обработка сигналов и её применение

Выпуск: XXVIII

Почтовый адрес:

107031, Москва, ул. Рождественка, д. 6/9/20

Исполнительная Дирекция РНТОРЭС им. А.С. Попова

E-mail организационного комитета конференции:
dspa@dspa-conf.org

Группа в Telegram:
https://t.me/+IFIWdpcBCO4yMTdi

Веб-сайт конференции: www.dspa-conf.org

Тел. РНТОРЭС им. Попова: +7 (903) 201-53-33