

DSPA:

Вопросы применения цифровой обработки сигналов

№3

2026

СОДЕРЖАНИЕ

Хайдаров А.Ф., Воронова Л.И. ЛЕГКОВЕСНЫЕ НЕЙРОСЕТЕВЫЕ АРХИТЕКТУРЫ ДЛЯ РАСПОЗНАВАНИЯ АКТИВНОСТИ ЧЕЛОВЕКА НА ПЕРИФЕРИЙНЫХ УСТРОЙСТВАХ	4
Широков В.А., Вовик А.Г. ИССЛЕДОВАНИЕ МОДЕЛЕЙ И МЕТОДОВ КЛАССИФИКАЦИИ И СЕГМЕНТАЦИИ ОЗЕР ПО СПУТНИКОВЫМ СНИМКАМ	11
Панков К.Н., Орлова А.С. ГОМОМОРФНОЕ ШИФРОВАНИЕ КАК МЕХАНИЗМ ЗАЩИТЫ ДАННЫХ В НЕДОВЕРЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СРЕДАХ	18
Бирюков Н.А., Синева И.С. РАЗРАБОТКА И ОПТИМИЗАЦИЯ ИКТ-РЕШЕНИЯ ДЛЯ ПАРНОГО ТРЕЙДИНГА: КЕЙС КОИНТЕГРИРОВАННОЙ ПАРЫ VISA/MASTERCARD	25
Фатхулин Т.Д., Метрик Е.А. АНАЛИЗ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ СПРОСА ПО ФИНАНСОВЫМ ПОКАЗАТЕЛЯМ	34
Федунов А.М., Кораблев Б.П., Якушин Д.А., Власюк И.В. МЕТОД АВТОМАТИЧЕСКОЙ ФОКУСИРОВКИ МАНУАЛЬНОГО ОБЪЕКТИВА НА ОСНОВЕ СТЕРЕОСКОПИЧЕСКОЙ ОЦЕНКИ ГЛУБИНЫ СЦЕНЫ	42

ЛЕГКОВЕСНЫЕ НЕЙРОСЕТЕВЫЕ АРХИТЕКТУРЫ ДЛЯ РАСПОЗНАВАНИЯ АКТИВНОСТИ ЧЕЛОВЕКА НА ПЕРИФЕРИЙНЫХ УСТРОЙСТВАХ

Хайдаров Адель Фанисович

Московский Технический Университет Связи и Информатики, магистрант, Москва, Россия
adelhaidar.se@gmail.com

Воронова Лилия Ивановна

Московский Технический Университет Связи и Информатики, д.ф.-м.н., проф., Москва, Россия
voronova.lilia@yandex.ru

Аннотация

В статье исследуются легковесные архитектуры нейронных сетей для распознавания активности человека (HAR) по данным инерциальных датчиков с ориентацией на развёртывание в системах Edge AI. Проведён сравнительный анализ шести архитектур, включая предложенную LiteHAR на основе inverted residual blocks с механизмом Squeeze-and-Excitation. Эксперименты на датасете UCI HAR показали, что LiteHAR достигает точности 95.8% при размере модели менее 1 МБ. Для устройств с жёсткими ограничениями памяти предложена ультралёгкая модель Light+Linear (93.8%, 49 КБ). Результаты работы применимы для создания интеллектуальных систем мониторинга в промышленности, здравоохранении и робототехнике.

Ключевые слова

распознавание активности человека, легковесные нейронные сети, инерциальные датчики, Edge AI, глубокое обучение, встраиваемые системы.

Введение

Распознавание активности человека (Human Activity Recognition, HAR) по данным инерциальных измерительных блоков (IMU) является ключевой задачей для носимых устройств и систем мониторинга здоровья [1]. Практические приложения HAR включают фитнес-трекинг, обнаружение падений у пожилых людей, реабилитационный мониторинг и контекстно-зависимые мобильные сервисы. Задача HAR также находит применение в промышленной робототехнике, где распознавание действий оператора позволяет реализовать интуитивное управление манипуляторами [2] и автоматизированными системами.

Классические подходы к HAR основывались на формировании признаков пространства экспертным путём: вычислении статистических дескрипторов (среднее, дисперсия, энтропия, коэффициенты асимметрии и эксцесса), частотных характеристик (коэффициенты дискретного преобразования Фурье, вейвлет-признаки, спектральная энтропия) и корреляций между каналами [3]. Такой подход требовал экспертных знаний в предметной области и существенно ограничивал переносимость решений на новые условия эксплуатации, типы датчиков и целевые платформы.

Современные методы глубокого обучения демонстрируют высокую точность на задачах HAR, однако требуют значительных вычислительных ресурсов [4]. Рекуррентные архитектуры (LSTM, GRU) эффективно моделируют временные зависимости, но имеют ограничения по параллелизации вычислений [5]. Свёрточные нейронные сети (CNN) обеспечивают иерархическое извлечение признаков и поддерживают параллельную обработку на современных ускорителях [6]. Transformer-архитектуры, успешно применяемые в обработке естественного языка и компьютерном зрении [7], также исследуются для задач анализа временных рядов.

Развёртывание моделей глубокого обучения на устройствах с ограниченными ресурсами – смартфонах, фитнес браслетах, микроконтроллерах – требует компактных архитектур с малым объёмом памяти (до 1–2 МБ) и низкой вычислительной сложностью [8]. Данное направление, известное как Edge AI или TinyML, приобретает особую актуальность в контексте развития Интернета вещей (IoT) и беспроводных сенсорных сетей [9, 10].

Методы оптимизации моделей для Edge AI включают несколько направлений: архитектурные инновации (depthwise-separable convolutions, inverted residual blocks, neural architecture search), методы

компрессии (квантизация весов, pruning, knowledge distillation) и использование специализированных аппаратных ускорителей (NPU, TPU). В данной работе исследуется первый подход – проектирование изначально эффективных архитектур без постобработки, что обеспечивает предсказуемое поведение модели на различных целевых платформах.

В рамках настоящего исследования проводится сравнительный анализ легких архитектур для HAR с количественной оценкой компромисса между точностью классификации, размером модели и относительной вычислительной сложностью. Результаты работы могут быть использованы при создании интеллектуальных систем управления, включая системы удалённого управления промышленным оборудованием и роботами [2], а также системы мониторинга состояния операторов технологических процессов [11].

Обзор методов распознавания активности

Задача распознавания активности человека по данным инерциальных датчиков активно исследуется в последние десятилетия. Ранние работы [1] предлагали комплексные конвейеры обработки данных, включающие сегментацию сигналов скользящим окном, извлечение признаков вручную и классификацию методами машинного обучения (SVM, Random Forest, k-NN).

Переход к методам глубокого обучения позволил автоматизировать процесс извлечения признаков. В работе [4] предложена архитектура DeepConvLSTM, сочетающая свёрточные слои для извлечения локальных паттернов с рекуррентными слоями для моделирования временных зависимостей. Данный подход достиг точности 91-93% на различных бенчмарках, однако высокая вычислительная сложность LSTM-слоёв ограничивает применимость на встраиваемых устройствах.

Архитектура MobileNet [12] продемонстрировала эффективность depthwise-separable свёрток для задач компьютерного зрения, сократив число параметров и вычислительную сложность в среднем в 8-9 раз по сравнению со стандартными свёртками при сопоставимой точности. Развитием данного подхода стала архитектура MobileNetV2 [13] с inverted residual blocks, где расширение признакового пространства выполняется в промежуточном слое, а не на входе блока.

Механизм Squeeze-and-Excitation (SE) [14] обеспечивает адаптивную перекалибровку каналов на основе глобальной контекстной информации. SE-блоки добавляют минимальное количество параметров (около 2%), но позволяют повысить точность на 0.5-1% за счёт явного моделирования взаимозависимостей между каналами. Архитектура EfficientNet [15] объединила данные подходы с методом compound scaling для сбалансированного масштабирования глубины, ширины и разрешения сети.

Архитектуры на основе механизма внимания (attention) [7] широко применяются для анализа последовательностей ввиду их способности эффективно извлекать зависимости между удалёнными элементами и возможности распараллеливания вычислений на больших выборках. Однако квадратичная сложность self-attention по длине последовательности ограничивает применимость на длинных временных рядах. Для коротких окон (128-256 отсчётов) Transformer-архитектуры демонстрируют конкурентоспособные результаты, но требуют большего числа параметров по сравнению с CNN.

Важным аспектом является интеграция моделей HAR в более сложные системы. Распознавание активности оператора может использоваться для адаптивного управления промышленными манипуляторами [2] и координации действий в гетерогенных сетях [9]. Моделирование маршрутизации в кластеризованных системах БПЛА [10] также требует учёта контекста человеческой активности для оптимизации энергопотребления узлов сети.

Данные и предобработка

В работе использован публичный датасет UCI Human Activity Recognition [3], собранный с участием 30 добровольцев (возраст 19-48 лет) со смартфоном Samsung Galaxy S II, закреплённым на поясе. Записи трёхосевого акселерометра и гироскопа соответствуют шести типам активности: ходьба по ровной поверхности (WALKING), подъём по лестнице (WALKING_UPSTAIRS), спуск (WALKING_DOWNSTAIRS), сидение (SITTING), стояние (STANDING) и лежание (LAYING).

Сырые сигналы предварительно фильтровались медианным фильтром для устранения импульсных помех и фильтром Баттерворта 3-го порядка с частотой среза 20 Гц для удаления высокочастотных шумов. Выбор частоты среза обусловлен частотным диапазоном человеческих движений (0.1-15 Гц для

большинства активностей). Затем сигналы сегментировались скользящим окном длительностью 2.56 с (128 отсчётов при частоте дискретизации 50 Гц) с перекрытием 50%. Перекрытие обеспечивает увеличение объёма обучающих данных и устойчивость к фазовым сдвигам.

Перед подачей в модель данные нормализовались поканалам: для каждого канала вычислялись среднее значение и стандартное отклонение на обучающей выборке, затем применялась стандартизация (z-score normalization). Нормализация выполняется независимо для каждого канала, что учитывает различия в диапазонах измерений акселерометра и гироскопа.

Входной тензор имеет размерность 9×128 : три канала `body_acc` (ускорение тела с вычтенной гравитационной составляющей, выделенной фильтром Баттерворта с частотой среза 0,3 Гц), три канала `body_gyro` (угловая скорость) и три канала `total_acc` (полное ускорение, включая гравитацию). Обучающая выборка содержит 7352 образца от 21 добровольца, тестовая – 2947 образцов от оставшихся 9 участников. Такое разбиение по участникам (subject-wise split) исключает утечку данных между выборками и обеспечивает оценку обобщающей способности модели на новых пользователях. Дополнительно 15% обучающих данных выделено для валидации.

Исследуемые архитектуры

Standard CNN

Базовая трёхуровневая свёрточная сеть служит референсной точкой для сравнения. Структура: Conv1d(9→64, k=7) → BN → ReLU → MaxPool(2) → Conv1d(64→128, k=5) → BN → ReLU → MaxPool(2) → Conv1d(128→256, k=3) → BN → ReLU → GAP → MLP(256→128→6). Здесь BN обозначает Batch Normalization для стабилизации распределения активаций, GAP – Global Average Pooling (глобальное усреднение по временной оси), MLP – многослойный перцептрон. Размер ядра $k=7$ на первом слое захватывает временные паттерны длительностью около 140 мс, что соответствует характерному времени одного шага при ходьбе. Уменьшающиеся размеры ядер ($7 \rightarrow 5 \rightarrow 3$) на последующих слоях позволяют извлекать всё более абстрактные признаки. Модель содержит 178К параметров (0.68 МБ в формате FP32).

Lightweight CNN

Стандартные свёртки заменены на depthwise-separable [12], состоящие из depthwise Conv (независимая фильтрация каждого канала с C группами) и pointwise Conv (1×1 для смешивания каналов). Это сокращает число параметров и операций в k раз, где k – размер ядра. Для ядра размером 7 теоретическое сокращение составляет $7\times$, на практике – около $6\times$ с учётом pointwise слоя. Структура: Conv(9→32) + $2 \times$ DSCConv + GAP + MLP. Размер модели – 29К параметров (0.11 МБ), что в 6 раз меньше базовой при потере точности менее 0.5%.

LiteHAR (предложенная)

Предложенная архитектура LiteHAR построена на Inverted Residual Blocks [13] с механизмом Squeeze-and-Excitation [14] (рис. 1). Данная комбинация позволяет эффективнее использовать параметры модели за счёт расширения признакового пространства в промежуточном слое и адаптивной перекалибровки каналов.

Inverted Residual Block состоит из трёх этапов: (1) Expansion – pointwise Conv (1×1) увеличивает число каналов в $t = 6$ раз, создавая высокоразмерное пространство признаков для последующей фильтрации; (2) Depthwise Conv 3×1 выполняет пространственную (временную) фильтрацию независимо для каждого канала; (3) Projection – pointwise Conv сжимает представление обратно к исходной размерности. Активация ReLU6 ($\min(\max(0, x), 6)$) ограничивает выходы диапазоном $[0, 6]$, что способствует стабильности при последующей квантизации модели.

SE Block выполняет адаптивную перекалибровку каналов на основе глобальной статистики. Веса важности каналов вычисляются как: $\mathbf{s} = \sigma(\mathbf{W}_2 \text{ReLU}(\mathbf{W}_1 \text{GAP}(\mathbf{x})))$, где $\mathbf{W}_1 \in \mathbb{R}^{C/r \times C}$, $\mathbf{W}_2 \in \mathbb{R}^{C \times C/r}$, коэффициент редукции $r = 4$, σ – сигмоидная функция. SE-блок применяет глобальное усреднение по временной оси (Global Average Pooling), затем два полносвязных слоя с редукцией размерности для вычисления весов важности каналов. Выход умножается поэлементно на входной тензор.

Структура сетки: Stem(9→32, s=2) → [InvRes(32)×2 + SE] → [InvRes(64, s=2)×2 + SE] → [InvRes(128, s=2)×2 + SE] → GAP → MLP(6).

Stem представляет собой начальный свёрточный слой Conv1d(9→32, k=3, s=2) с Batch Normalization и активацией ReLU6, выполняющий предварительную обработку входных данных и

уменьшение временного разрешения в два раза. Последовательное уменьшение разрешения (stride=2) на каждом этапе позволяет увеличивать рецептивное поле при фиксированном размере ядра. GAP преобразует тензор размерности (B, C, T) в (B, C), устраняя зависимость от длины входной последовательности. Модель содержит 255K параметров (0.97 МБ).

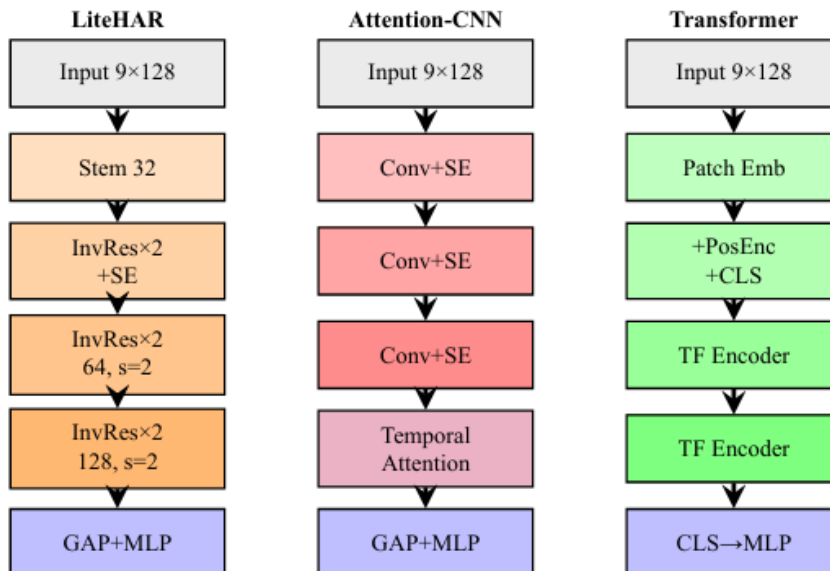


Рис. 1. Архитектуры: LiteHAR, Attention-CNN и Transformer

Inverted residuals эффективнее классических bottleneck за счёт расширения в промежуточном слое, где depthwise conv работает с большим числом каналов. SE-блоки дают прирост точности около 0.3% при росте числа параметров на 2%.

Attention-CNN

Архитектура сочетает SE-блоки после каждого свёрточного слоя с модулем Temporal Self-Attention (4 головы) для моделирования долгосрочных зависимостей во временном ряде. Multi-head attention позволяет модели одновременно обращать внимание на различные временные масштабы, что важно для активностей с периодической структурой (ходьба, бег). Число параметров – 486K (1.85 МБ).

Transformer

Входной сигнал преобразуется в последовательность патчей через линейную проекцию Conv1d(9→128, k=5). К эмбедингам добавляется обучаемое позиционное кодирование и специальный CLS-токен для агрегации информации. Два слоя Transformer Encoder (4 головы, FFN dim=512, dropout=0.1) обрабатывают последовательность, классификация выполняется по выходу CLS-токена через MLP. Число параметров – 437K (1.67 МБ).

Light+Linear

Ультралёгкая модель для устройств с жёсткими ограничениями ресурсов, таких как микроконтроллеры класса ARM Cortex-M4/M7 [16]. Структура аналогична Lightweight CNN, но MLP-голова заменена единственным линейным слоем без скрытых нейронов: Conv(9→32) + 2×DSCConv + GAP + Linear(6). Отсутствие нелинейного классификатора компенсируется нелинейностями в свёрточной части сети. Число параметров – 13K (49 КБ). Модель потенциально пригодна для развёртывания с использованием TensorFlow Lite Micro или ONNX Runtime Mobile [17].

Условия эксперимента

Обучение проводилось с оптимизатором AdamW ($\beta_1=0.9$, $\beta_2=0.999$, weight decay 10^{-4}), начальной скоростью обучения 10^{-3} и расписанием CosineAnnealing с тёплым перезапуском. Применялись label smoothing (коэффициент 0.1) для регуляризации и предотвращения переобучения на шумных метках,

а также gradient clipping (порог 1.0) для стабилизации обучения при больших градиентах. Размер батча – 64 образца.

Для аугментации данных использовались три преобразования, применяемые с вероятностью 0.5 каждое: аддитивный гауссов шум ($\sigma=0.02$) для повышения робастности к сенсорному шуму, случайный циклический сдвиг по времени (± 10 отсчётов) для инвариантности к фазе сигнала и масштабирование амплитуды сигнала (коэффициент 0.95-1.05) для устойчивости к калибровочным погрешностям датчиков. Аугментация применялась только к обучающей выборке.

Обучение останавливалось при отсутствии улучшения на валидационной выборке в течение 10 эпох (early stopping). Максимальное число эпох – 100. Каждая конфигурация запускалась 6 раз с различными seed.

Время инференса измерялось на CPU как среднее значение по 1000 последовательных запусков одного батча размером 64 образца. Данные измерения предназначены для *относительного сравнения* вычислительной сложности архитектур и не отражают реальную производительность на целевых Edge-устройствах. Для оценки производительности на конкретной целевой платформе требуется дополнительное профилирование с учётом особенностей аппаратного обеспечения, оптимизаций компилятора и используемого фреймворка инференса.

Размер модели указан для формата FP32 (32-битное представление с плавающей точкой), что соответствует стандартному формату весов PyTorch без квантизации.

Результаты и обсуждение

Сводные результаты экспериментов приведены в табл. 1. LiteHAR показала наивысшую точность 95.8% (F1=95.8%), опередив Attention-CNN (95.4%) и Transformer (95,0%) при меньшем размере модели. Превосходство LiteHAR над Attention-CNN объясняется более эффективным использованием параметров: inverted residuals создают оптимальное промежуточное представление при минимальном числе весов, тогда как attention-слои требуют дополнительных проекционных матриц.

Таблица 1

Сравнение архитектур на UCI HAR (по 6 запускам)

Модель	Acc, %	F1, %	Par., K	Size, MB	Inf*, ms
Standard	93.3±0.14	93.5	178	0.68	0.75
Lightweight	93.8±0.08	93.9	29	0.11	0.93
LiteHAR	95.8±0.12	95.8	255	0.97	4.18
Attention-CNN	95.4±0.07	95.4	486	1.85	2.20
Transformer	95.0±0.32	94.9	437	1.67	2.06
Light+Linear	93.8±0.24	93.8	13	0.05	0.57

Стандартное отклонение точности для LiteHAR ($\pm 0.12\%$) сопоставимо с другими CNN-архитектурами, что свидетельствует о стабильности обучения. Transformer демонстрирует наибольшую вариативность ($\pm 0.32\%$) вследствие чувствительности механизма внимания к инициализации весов.

Ультралёгкая Light+Linear достигла 93.8% при размере 49 КБ. Потеря 2% точности по сравнению с LiteHAR компенсируется 20-кратным сокращением размера модели, что критично для устройств с памятью менее 100 КБ. Данная архитектура рекомендуется для сценариев с жёсткими ограничениями ресурсов.

Анализ матрицы ошибок LiteHAR (табл. 2) показывает, что динамические активности распознаются с F1 выше 98% (ходьба – 99.8%, подъём – 98.4%, спуск – 98.5%), тогда как статические – на уровне 87-91%. Основной источник ошибок – путаница между SITTING и STANDING (90 случаев из 491 для SITTING): в обоих состояниях IMU регистрирует близкие к нулю значения ускорений и угловых скоростей, различие определяется преимущественно ориентацией устройства относительно вектора гравитации.

Таблица 2

Матрица ошибок LiteHAR и F1-score по классам

	WK	UP	DN	SIT	STD	LAY	F1
WK	494	0	2	0	0	0	99.8
UP	2	461	8	0	0	0	98.4
DN	0	0	420	0	0	0	98.5
SIT	0	2	0	394	90	5	87.6
STD	0	0	0	20	512	0	90.6
LAY	0	0	0	0	0	537	99.7

Полученные результаты согласуются с опубликованными работами на данном датасете. Достигнутая точность 95.8% превосходит результаты DeepConvLSTM [4] (92–93%) и сопоставима с лучшими опубликованными результатами при существенно меньшем размере модели.

Заключение

В работе исследованы легковесные архитектуры нейронных сетей для распознавания активности человека по данным инерциальных датчиков. Предложенная архитектура LiteHAR на основе inverted residual blocks и механизма Squeeze-and-Excitation достигает точности 95.8% на датасете UCI HAR при размере менее 1 МБ. Для устройств с жёсткими ограничениями памяти рекомендуется модель Light+Linear (93.8%, 49 КБ), пригодная для развёртывания на микроконтроллерах.

Экспериментально подтверждено, что inverted residuals превосходят классические свёртки благодаря расширению признакового пространства в промежуточном слое, где depthwise convolution работает с большим числом каналов. SE-блоки обеспечивают дополнительный прирост точности при минимальном увеличении числа параметров за счёт адаптивной перекалибровки каналов. Transformer-архитектура не даёт преимуществ на коротких временных окнах (2.56 с) ввиду ограниченной длины контекста и избыточности механизма self-attention для данной задачи.

Результаты работы могут быть использованы при создании интеллектуальных систем мониторинга активности в здравоохранении, промышленности и робототехнике. Интеграция моделей HAR с системами управления промышленным оборудованием [2] и координации в беспроводных сенсорных сетях [9, 10] закладывает основу для создания адаптивных человеко-машинных интерфейсов.

Направления дальнейших исследований включают квантизацию моделей (INT8/INT4), валидацию на датасетах WISDM [18] и PAMAP2 [19] для оценки обобщающей способности, профилирование на целевых встраиваемых платформах (ARM Cortex-M, RISC-V), а также интеграцию с цифровыми двойниками технологических процессов [11] для предиктивного мониторинга операторов.

Литература

1. Bulling A., Blanke U., Schiele B. A tutorial on human activity recognition using body-worn inertial sensors // ACM Computing Surveys. 2014. Vol. 46, № 3, pp. 1-33.
2. Белов Н.В., Воронова Л.И. Система удалённого управления промышленным манипулятором KUKA // Автоматизация в промышленности. 2023. № 12. DOI: 10.25728/avtprom.2023.12.09.
3. Anguita D. et al. A public domain dataset for human activity recognition using smartphones // Proc. European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN). Bruges, 2013, pp. 437-442.
4. Ordóñez F.J., Roggen D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition // Sensors. 2016. Vol. 16, № 1. Art. 115.
5. Hochreiter S., Schmidhuber J. Long short-term memory // Neural Computation. 1997. Vol. 9, № 8, pp. 1735-1780.
6. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. 2015. Vol. 521, № 7553, pp. 436-444.
7. Vaswani A. et al. Attention is all you need // Advances in Neural Information Processing Systems (NeurIPS). 2017. Vol. 30, pp. 5998–6008.
8. Banbury C. et al. Benchmarking TinyML systems: Challenges and direction // arXiv preprint arXiv:2003.04821. 2020.

9. *Mohammad N.F., Voronova L.I., Voronov V.I., Rozhkov S.A.* Software complex for modelling routing in heterogeneous model of wireless sensor network // Proc. IEEE Systems of Signals Generating and Processing in the Field of on Board Communications. Moscow, 2024. P. 1-5. DOI: 10.1109/IEEECONF60226.2024.10496736.
10. *Мохаммад Н., Воронова Л.И., Воронов В.И.* Программа для моделирования маршрутизации в кластеризованном роуе БПЛА с использованием генетического алгоритма // Первая миля. 2023. DOI: 10.22184/2070-8963.2023.114.6.46.52.
11. *Smolnikov V.A., Voronova L.I., Voronov V.I., Rozhkov S.A., Petukhov V.M.* Simulation of the digital twin of the technological process of creating a demonstrator using R-PRO Digital // Proc. IEEE Systems of Signals Generating and Processing in the Field of on Board Communications. Moscow, 2024. P. 1-5. DOI: 10.1109/IEEECONF60226.2024.10496776.
12. *Howard A.G. et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications // arXiv preprint arXiv:1704.04861. 2017.
13. *Sandler M. et al.* MobileNetV2: Inverted residuals and linear bottlenecks // Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, 2018, pp. 4510-4520.
14. *Hu J., Shen L., Sun G.* Squeeze-and-Excitation networks // Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, 2018, pp. 7132-7141.
15. *Tan M., Le Q.V.* EfficientNet: Rethinking model scaling for convolutional neural networks // Proc. 36th International Conference on Machine Learning (ICML). Long Beach, 2019. Vol. 97, pp. 6105-6114.
16. *David R. et al.* TensorFlow Lite Micro: Embedded machine learning for TinyML systems // Proc. Machine Learning and Systems (MLSys). 2021. Vol. 3, pp. 800-811.
17. ONNX Runtime [Электронный ресурс]. URL: <https://onnxruntime.ai/> (дата обращения: 15.01.2026).
18. *Kwapisz J.R., Weiss G.M., Moore S.A.* Activity recognition using cell phone accelerometers // ACM SIGKDD Explorations Newsletter. 2011. Vol. 12, № 2, pp. 74-82.
19. *Reiss A., Stricker D.* Introducing a new benchmarked dataset for activity monitoring // Proc. 16th International Symposium on Wearable Computers (ISWC). Newcastle, 2012, pp. 108-109.

ИССЛЕДОВАНИЕ МОДЕЛЕЙ И МЕТОДОВ КЛАССИФИКАЦИИ И СЕГМЕНТАЦИИ ОЗЕР ПО СПУТНИКОВЫМ СНИМКАМ

Широков Владимир Андреевич
МТУСИ, студент, Москва, Россия
vovashirokov003@gmail.com

Вовик Андрей Геннадьевич
МТУСИ, старший преподаватель, к.т.н. Москва, Россия
a.g.vovik@mtuci.ru

Аннотация

В статье проведен анализ методов классификации и сегментации озёр по данным спутниковых снимков. Проанализированы основные подходы к автоматизированной сегментации водных объектов, включая пороговые методы, индексные показатели и алгоритмы машинного обучения. Оценены преимущества и ограничения различных методов с точки зрения точности, устойчивости к помехам и применимости к снимкам разного пространственного разрешения.

Ключевые слова

спутниковые снимки, классификация и сегментация, озера, машинное обучение, гидрология

Введение

Озёра являются важнейшими компонентами природных экосистем и играют значительную роль в формировании гидрологического режима, биологического разнообразия и климатических условий территорий. Изменения площади и конфигурации озёр могут служить индикаторами климатических колебаний, антропогенного воздействия и деградации природных ландшафтов. В связи с этим актуальной является задача оперативного и точного мониторинга водных объектов.

Современные методы дистанционного зондирования Земли позволяют получать регулярные спутниковые данные с высоким пространственным и временным разрешением, что создаёт широкие возможности для исследования динамики озёрных систем. Однако обработка больших объёмов спутниковых изображений требует применения автоматизированных методов классификации и сегментации, обеспечивающих выделение водных объектов с минимальным участием оператора [1].

На сегодняшний день разработано множество подходов к выделению водных поверхностей на спутниковых снимках, включая использование спектральных водных индексов [2], пороговых алгоритмов [3], методов классификации на основе признаков [4], а также алгоритмов машинного и глубокого обучения [5, 6]. Каждый из этих методов обладает своими преимуществами и ограничениями, связанными с типом используемых данных, условиями съёмки и особенностями исследуемой территории.

Целью данной работы является исследование методов классификации и сегментации озёр по спутниковым снимкам. В рамках исследования рассматриваются различные алгоритмы выделения водных объектов и оценивается их эффективность при обработке спутниковых данных. Полученные результаты могут быть использованы при решении задач мониторинга водных ресурсов и анализа пространственно-временной динамики озёр.

Обоснование выбора методов классификации и сегментации

Автоматизированное выделение и типологическая классификация озёр по спутниковым данным представляет собой комплексную задачу, включающую этапы детекции водных объектов, их пространственной сегментации и последующего отнесения к определённым генетическим типам. Земля и машинного обучения показывает, что эффективность решения данной задачи в значительной степени определяется корректным выбором алгоритмов обработки изображений и моделей классификации [4].

1. Анализ подходов к сегментации водных объектов

На первом этапе обработки спутниковых изображений необходимо выполнить сегментацию - выделение пикселей, соответствующих водной поверхности. В научной литературе можно выделить три основных группы методов сегментации водных объектов:

1.1. Индексные методы

Одним из наиболее распространённых и базовых подходов является использование спектральных водных индексов, таких как NDWI, MNDWI и AWEI. Данные методы основаны на различиях отражательной способности воды и суши в видимом и ближнем инфракрасном диапазонах электромагнитного спектра, что позволяет эффективно отделять водную поверхность от окружающих территорий [7]. Простота реализации и низкая вычислительная сложность делают индексные методы удобным инструментом для обработки больших массивов спутниковых данных, в том числе снимков Sentinel-2 и Landsat, при отсутствии обучающей выборки.

При благоприятных условиях съёмки спектральные индексы обеспечивают высокую точность бинарного разделения классов «вода – не вода». Вместе с тем их применение существенно ограничено в сложных ландшафтных условиях. Наличие теней, тёмных почв, прибрежной растительности, а также льда, взвешенных частиц и мелководных участков приводит к зашумлению спектрального сигнала и снижению качества сегментации. Кроме того, индексные методы не позволяют различать типы водных объектов, поскольку основаны исключительно на спектральных признаках и не учитывают морфологические и пространственные характеристики озёр.

Анализ работ, посвящённых дистанционному мониторингу водных ресурсов, показывает, что спектральные индексы целесообразно рассматривать в качестве инструмента первичной сегментации водной поверхности. Однако для задач, связанных с детальным анализом формы береговой линии и типологической классификацией озёр, их применение является недостаточным и требует использования более сложных методов обработки спутниковых изображений [5]. В таблице 1 приведены оценки точности выделения водных объектов для различных водных индексов.

Таблица 1

Точность выделения водных объектов для водных индексов

Название метода	Общая точность, %
NDVI	≈ 95,85
AWEI	≈ 99,00
NDWI	≈ 98,40
NDMI	≈ 98,05
MNDWI	≈ 98,55
WRI	≈ 98,70




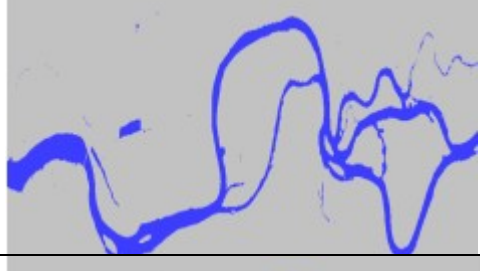


В таблице 2 приведена визуализация полученных «масок» водных объектов, полученных с использованием разных индексов

Представленные данные свидетельствуют о высокой общей точности всех рассмотренных спектральных индексов при идентификации водных объектов по материалам дистанционного зондирования. Максимальное значение точности характерно для индекса AWEI, что указывает на его наибольшую эффективность и устойчивость к фоновым искажениям. Минимальная точность наблюдается у NDVI, что обусловлено его преимущественной ориентированностью на оценку растительного покрова. В целом различия между специализированными водными индексами носят незначительный характер, что позволяет рассматривать их как надёжные инструменты для решения задач выделения водных объектов.

Сравнивая результаты, представленные в таблицах 1 и 2, можно отметить, что наиболее хорошо выделяется водный объект с помощью индекса AWEI [8].

Таблица 2

Визуализация маски водных объектов

Индекс	Изображение маски водных объектов
NDVI	
AWEI	
NDWI	
NDMI	
MNDWI	
WRI	

1.2. Классические методы машинного обучения

К данной группе относятся методы машинного обучения, такие как SVM, k-means, случайный лес и алгоритмы градиентного бустинга, которые применяются как в пиксельной, так и в объектно-ориентированной сегментации спутниковых изображений. Их использование позволяет учитывать совокупность разнородных признаков, включая спектральные характеристики, текстурные параметры и геометрические особенности объектов, что обеспечивает более высокую точность сегментации по сравнению с индексными подходами. Особенностью алгоритмов ансамблевого типа, в частности случайного леса, является относительная интерпретируемость результатов, что облегчает анализ вклада отдельных признаков в процесс классификации.

Вместе с тем применение классических методов машинного обучения связано с рядом ограничений. Для их эффективной работы требуется предварительное формирование признакового пространства, что предполагает ручной отбор информативных параметров и может существенно влиять на итоговое качество сегментации. Кроме того, данные методы в ограниченной степени учитывают пространственный контекст и взаимное расположение пикселей, что снижает их устойчивость при обработке водных объектов со сложной геометрией береговой линии и высокой морфологической вариативностью. Как показывают результаты исследований, методы случайного леса и градиентного бустинга обеспечивают стабильные показатели при сегментации водных объектов, однако чувствительны к масштабной изменчивости форм озёр и качеству выбранных признаков [4].

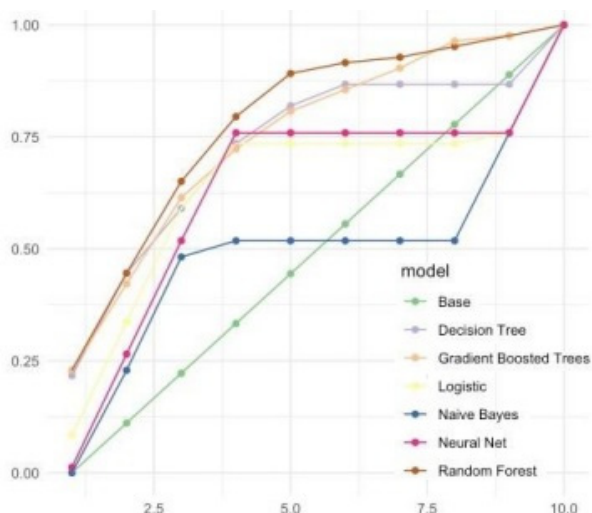


Рис. 1. Точность разных моделей

1.3. Модели глубокого обучения

Современные работы в области обработки спутниковых изображений показывают, что для задач семантической сегментации водных объектов наиболее эффективными являются методы глубокого обучения, основанные на сверточных нейронных сетях. Наибольшее распространение получили архитектуры U-Net, DeepLab и SegNet, обеспечивающие высокую точность выделения водной поверхности за счёт совместного учёта спектральной информации и пространственного контекста изображения [9]. В отличие от классических алгоритмов, данные модели автоматически формируют набор информативных признаков, что позволяет корректно восстанавливать форму береговой линии и снижать влияние шумов, неоднородного освещения и фоновых объектов.

В то же время применение нейросетевых методов сопряжено с рядом ограничений, среди которых ключевыми являются высокая вычислительная нагрузка и необходимость использования значительных объёмов размеченных данных для обучения моделей. Кроме того, внутренние представления, формируемые сверточными нейронными сетями, остаются слабо интерпретируемыми, что усложняет анализ причин ошибок сегментации и оценку вклада отдельных признаков.

С учётом указанных факторов в качестве основной технологической основы настоящего исследования были выбраны архитектуры U-Net и DeepLab, как наиболее универсальные и адаптированные к работе с многоспектральными спутниковыми данными. Данный выбор опирается на результаты, представленные в работе «Deep learning-based multi-spectral satellite image segmentation for water

body detection» [10], где показана высокая эффективность указанных архитектур при выделении водных объектов в условиях различного пространственного разрешения и спектрального состава данных.

Выбор типов озёр для классификации по спутниковым снимкам

Одним из ключевых этапов исследования, направленного на автоматизированную классификацию озёр по спутниковым снимкам, является обоснованный выбор типов озёр, подлежащих классификации. В классических лимнологических работах, в том числе в монографии Н. В. Мякишевой «Многокритериальная классификация озёр» [11], подчёркивается, что типологическое разнообразие озёр чрезвычайно велико и определяется совокупностью генетических, морфометрических, гидрологических и физико-географических факторов. В рамках многокритериального подхода классификация озёр предполагает учёт большого числа параметров, включая происхождение котловины, режим водообмена, химический состав воды, климатические условия и динамику уровня, что позволяет формировать устойчивые и научно обоснованные классы водоёмов.

В то же время при использовании спутниковых данных возможности наблюдения существенно ограничены набором признаков, доступных дистанционно. Космические снимки в видимом и ближнем инфракрасном диапазонах позволяют надёжно фиксировать лишь часть характеристик озёр, прежде всего их пространственное положение, площадь, форму котловины, конфигурацию береговой линии и особенности окружающего ландшафта. Такие параметры, как глубина, гидрологический режим или химический состав воды, в большинстве случаев не поддаются прямому определению по данным дистанционного зондирования. В связи с этим использование полной многокритериальной классификации в её классическом виде оказывается затруднительным для задач автоматизированной обработки спутниковых изображений.

С учётом изложенного в настоящем исследовании реализуется адаптация многокритериального подхода к условиям анализа космических данных, при которой в качестве базовых классификационных признаков рассматриваются морфометрические и пространственные характеристики озёр, устойчиво выявляемые по спутниковым снимкам. Анализ работ по озероведению и дистанционному зондированию позволяет утверждать, что наибольшую перспективу для автоматизированной классификации представляют типы озёр, обладающие выраженными и относительно стабильными геометрическими признаками, обусловленными их генезисом.

В рамках исследования были выбраны следующие основные типы озёр: тектонические, ледниковые, старичные, вулканические, карстовые и термокарстовые. Данные группы широко представлены в существующих типологических схемах и при этом отличаются морфологическими особенностями, которые могут быть количественно описаны на основе спутниковых изображений. Тектонические озёра, как правило, характеризуются крупными размерами и вытянутой формой котловины, что отражает структурно-тектоническую природу их происхождения. Эти особенности позволяют достаточно надёжно идентифицировать такие объекты на космических снимках по соотношению длины и ширины, компактности формы и ориентации относительно региональных тектонических структур [9, 11].


Ледниковые озёра отличаются более правильной формой котловин, часто близкой к округлой или овальной, а также характерным размещением в районах древнего или современного оледенения. Их морфология и пространственный контекст, включая плотность распределения и связь с рельефом, делают возможным выделение данного типа озёр на основе сочетания геометрических и ландшафтных признаков. Вулканические озёра, формирующиеся в кратерах и кальдерах, обладают замкнутыми, относительно симметричными формами и чётко выраженными границами, что обеспечивает их высокую различимость на спутниковых снимках даже при среднем пространственном разрешении.

Карстовые озёра характеризуются более сложной конфигурацией береговой линии и часто приурочены к районам развития растворимых пород. Несмотря на вариативность формы, их пространственная локализация и морфологические особенности позволяют рассматривать данный тип как потенциально различимый при использовании объектно-ориентированных методов анализа изображений. Особое место в исследовании занимают термокарстовые озёра, широко распространённые в криолитозоне. Их неправильная форма, динамика площади и специфическое пространственное распределение в пределах зон многолетнемёрзлых пород создают набор признаков, который может быть использован для автоматизированной классификации при наличии временных рядов спутниковых данных.

Выбор указанных типов озёр обусловлен тем, что они удовлетворяют ключевым требованиям, предъявляемым к объектам классификации в задачах машинного и глубокого обучения [7]: наличие устойчивых морфометрических признаков, воспроизводимости на данных различного пространственного разрешения и возможности формального описания в виде числовых характеристик. Такой подход позволяет, с одной стороны, сохранить связь с классическими лимнологическими концепциями, а с другой – обеспечить практическую реализуемость автоматизированной классификации по спутниковым снимкам [12].

Таблица 3

Классификация озёр

Тип озера	Вид со спутников
Тектоническое	
Ледниковые	
Старичные	
Вулканические	
Карстовые	
Термокарстовые	

Заключение

В ходе работы рассмотрены методы сегментации и классификации озёр по данным дистанционного зондирования Земли с целью оценки их применимости в задачах автоматизированного мониторинга водных объектов. Анализ классических лимнологических и географических исследований показал, что традиционные многокритериальные классификации озёр, основанные на учёте генезиса, морфометрических характеристик и физико-географических условий формирования, обладают ограниченной применимостью при автоматической обработке спутниковых данных. Как отмечается в работах Н. В. Мякишевой, подобные классификационные схемы ориентированы на комплексный анализ разнородных признаков и допускают неопределённость исходной информации, что затрудняет их формализацию в алгоритмическом виде [11].

Установлено, что при использовании спутниковых снимков целесообразно ограничивать типологическую классификацию теми типами озёр, которые характеризуются устойчивыми морфометрическими и пространственными признаками. В связи с этим в работе обоснован выбор тектонических, ледниковых, вулканических, карстовых и термокарстовых озёр как наиболее перспективных объектов автоматизированной классификации.

Показано, что спектральные индексные методы обеспечивают эффективное первичное выделение водной поверхности, однако не позволяют в полной мере решать задачи типологической дифференциации. В результате установлено, что наибольшую практическую эффективность демонстрирует комбинированный подход, включающий спектральную сегментацию и последующую классификацию сегментированных объектов с использованием методов машинного обучения, в частности ансамблевых алгоритмов, отличающихся устойчивостью и интерпретируемостью.

Исходя из вышперечисленного был сделан следующий вывод применение комбинированного подхода позволяет обеспечить оптимальное соотношение точности, вычислительной сложности и практической применимости при автоматизированной классификации озёр по спутниковым снимкам и может быть использовано при решении задач мониторинга и анализа динамики озёрных систем.

Литература

1. Шиверский Г. В., Кривошеков С. Н. Перспективы применения методов искусственного интеллекта в нефтегазовой геологии // MASTER'S JOURNAL. 2025. Том 28. № 4.
2. Канунова Е. Е., Садыков С. С. Алгоритмы пороговой сегментации для устранения дефектов на изображениях архивных документов. Муромский институт (филиал) Владимирского государственного университета. 2005. 7 с.
3. Ген А. С., Сорокин А. А., Шестаков Н. В. Бинарная классификация временных рядов полного электронного содержания методами машинного обучения на основе автоматических извлечённых признаков // VIII Международной конференции по глубокому обучению в вычислительной физике. Москва, 19-21 июня.
4. Сейдаметова З. С. Алгоритмы машинного и глубокого обучения // Информационно-компьютерные технологии в экономике, образовании и социальной сфере. 2019. № 4 (26). С. 5-13.
5. Катаев М. Ю., Бекеров А. А. Методика обнаружения водных объектов по многоспектральным спутниковым измерениям // Доклады Томского государственного университета систем управления и радиоэлектроники. 2017. -Т. 20, № 4. С. 105-108.
6. Емельянов А. А., Вовик А. Г., Лобаев А. А. Интеллектуальный анализ изменения водных ресурсов с использованием спутниковых снимков sentinel-2 // Телекоммуникации и информационные технологии. 2025. Т. 12, № 1. С. 90-96. EDN RKFHRJ
7. Терехин Э. А. Методика поиска эффективных спектральных индексов для автоматизированного дешифрирования водных объектов (на примере Белгородской области) // География и природные ресурсы. 2013. № 3. С. 132-138.
8. Kunhao Yuan, Xu Zhuang, Gerald Schaefer, Jianxin Feng, Lin Guan, Hui Fang. Deep-Learning-Based Multi-spectral Satellite Image Segmentation for Water Body Detection.
9. Swapna J., Khalid M. M. Remote Sensing Techniques for Water Quality Monitoring: A Review. Sensors (Basel). 2024. Vol. 24, № 24. Article 8041. DOI:10.3390/s24248041.
10. Фролов А. И. Анализ алгоритма градиентного бустинга для целей прогнозирования // Альманах научных работ молодых учёных: XLVIII науч. и учебно-метод. конф. Университета ИТМО. Том 1. С. 285-288.
11. Мякишева Н.В. Многокритериальная классификация озёр. СПб.: изд. РГГМУ, 2009. 160 с.
12. Jawak S., Kulkarni K, Luis A. A Review on Extraction of Lakes from Remotely Sensed Optical Satellite Data with a Special Focus on Cryospheric Lakes. Advances in Remote Sensing, 2015. 4, pp. 196-213. doi: 10.4236/ars.2015.43016.

ГОМОМОРФНОЕ ШИФРОВАНИЕ КАК МЕХАНИЗМ ЗАЩИТЫ ДАННЫХ В НЕДОВЕРЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СРЕДАХ

Панков Константин Николаевич

*МТУСИ, заведующий кафедрой «Теория вероятностей и прикладная математика»,
к.ф.-м.н., доцент, Москва, Россия*
pankov_kn@mtuci.ru

Орлова Александра Сергеевна

МТУСИ, студент, Москва, Россия
orlovaalex1612@gmail.com

Аннотация

В статье рассматривается гомоморфное шифрование как механизм обеспечения конфиденциальности данных при обработке в недоверенных вычислительных средах. Описаны базовые принципы работы и основные виды гомоморфных схем, а также типовые модели угроз, связанные с обработкой данных на стороне облачной или внешней инфраструктуры. Отдельное внимание уделено границам применимости гомоморфного шифрования и необходимости его использования совместно с управлением ключами и контролем доступа. Рассмотрена роль Альянса HES при HomomorphicEncryption.org, разрабатывающего рекомендации по безопасности, параметрам и совместимости внедрений.

Ключевые слова

информационная безопасность, гомоморфное шифрование; полностью гомоморфное шифрование; частично гомоморфное шифрование; недоверенная среда; конфиденциальные вычисления; облачные вычисления; модели угроз; стандартизация; HES; HomomorphicEncryption.org.

Введение

Сегодня данные всё чаще обрабатываются не на личном компьютере, а в облаке, у подрядчика или в распределённой инфраструктуре. Из этого следует актуальность изучения задачи защиты данных при их обработке [1]. При этом важно защитить не только хранение и передачу, но и сам момент вычислений – именно там чаще всего появляются утечки.

Даже если вычислительная (серверная) сторона честно выполняет поставленную задачу, она может быть излишне любопытна, скомпрометирована или просто быть плохо настроенной. А значит, обработка данных в открытом виде на стороне сервера становится слабым местом, возникает проблема доверия к ней.

Гомоморфное шифрование (далее – ГШ) закрывает “дыру” конфиденциальности данных в использовании [2]: сервер может использовать данные в вычислениях, не видя их в открытом виде. Это особенно полезно там, где доверие к вычислительной среде ограничено.

Целью данной работы является рассмотрение ГШ как механизма обеспечения конфиденциальности данных при их обработке в недоверенных вычислительных средах, описание базовых принципов работы и основных видов гомоморфных схем, а также описание типовых моделей угроз, связанных с обработкой данных на стороне облачной или внешней инфраструктуры.

Поставленные задачи лежат в сфере изучения информационной безопасности существующих информационных систем с точки зрения обеспечения конфиденциальности криптографическими методами, продолжая серию работ [3-7] и, наконец [8]. Отметим, что изучение ГШ тесно связано с математическими проблемами теории защиты информации, которые необходимо решать для построения стойких к различным атакам криптографических систем [9-21].

Понятие и базовые принципы гомоморфного шифрования

Сперва дадим необходимое определение. Гомоморфное шифрование, в соответствии с ГОСТ Р 34.14-2025, это шифрование, при котором шифртекст для результата некоторой алгебраической опе-

рации над открытыми текстами можно построить, применив (возможно, другую) алгебраическую операцию к соответствующим им шифртекстам, без знания криптографического ключа.

При классификации на практике, обычно, выделяют три уровня схем ГШ: частично гомоморфные схемы поддерживают один тип операций (например, только сложение); несколько (ограниченно) гомоморфные (неполные) — и сложение, и умножение, но с ограничением по глубине вычислений; полностью гомоморфные (далее – ПГШ или FHE в англоязычном сокращении) позволяют выполнять произвольные вычисления, но обычно требуют больше ресурсов [22].

Опишем общую модель работы ГШ. Типовой сценарий выглядит так: владелец данных зашифровывает их у себя, сервер выполняет вычисления над шифртекстами, а расшифровать итог может только владелец ключа. На рисунке 1 показана общая логика процесса.

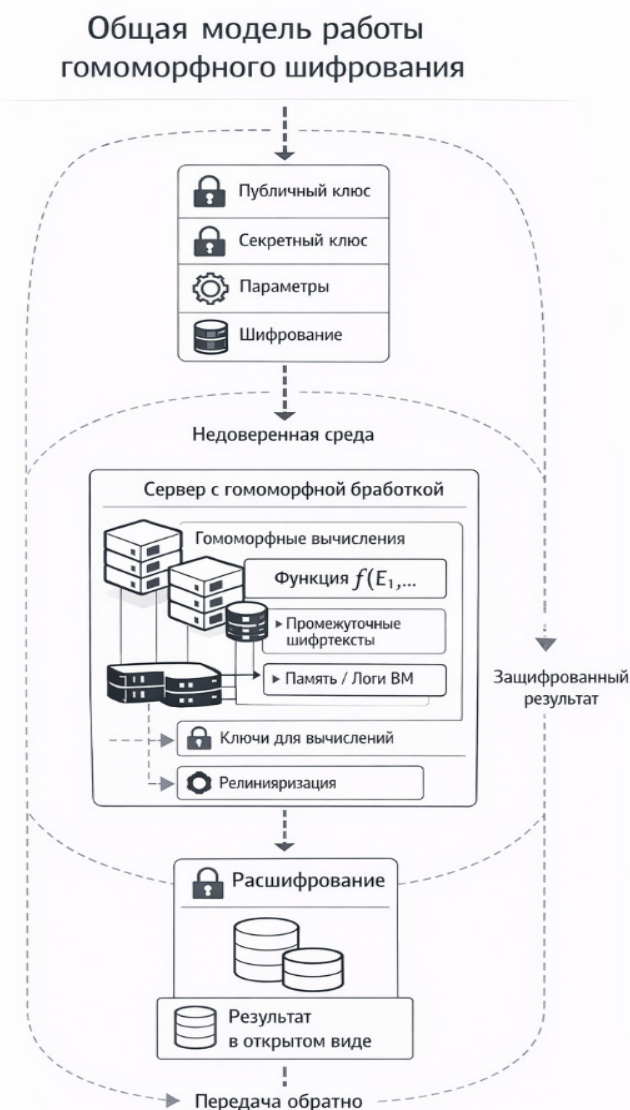


Рис. 1. Общая модель работы гомоморфного шифрования (шифрование – вычисление – расшифрование)

История развития гомоморфного шифрования

Идея “вычислений по зашифрованному” обсуждалась десятилетиями. Она тесно взаимосвязана с проблемой обеспечения безопасности облачных вычислений, идея которых появилась еще в 1960 г., когда Джон Маккарти высказал предположение, что когда-нибудь компьютерные вычисления будут производиться с помощью «общенародных утилит».

В 1978 г. в работе [23] впервые обосновали, что методом, позволяющим успешно проводить операции над зашифрованными данными, не искажая и не расшифровывая их, является то, что сейчас называется ГШ. В этой работе была описана концепция ГШ, а также авторы задались вопросами, возможно ли такое шифрование в принципе и для каких алгебраических систем такой гомоморфизм существует. Долгое время после этой работы на практике применялись в основном частично гомоморфные схемы – потому что это было реалистично по скорости и памяти.

Первая система ПГШ была представлена в 2009 году в диссертационной работе Крейга Джентри [24]. Появление ПГШ стало важным шагом: стало понятно, что вычислять по зашифрованному можно не только “чуть-чуть”, а в принципе для произвольных функций. Однако первые подходы были тяжёлыми и плохо подходили для повседневных систем.

Со временем схемы и реализации ускорялись, улучшались параметры безопасности и инструменты разработчика. Иначе говоря, происходила эволюция ГШ с точки зрения практического применения. Сегодня ГШ всё чаще рассматривают как прикладной механизм для задач аналитики, совместных вычислений и приватной обработки данных.

В настоящее время Одна из ключевых тенденций в развитии ГШ – это стандартизация [25-28] и согласование практик внедрения. Здесь важную роль играет Альянс HES (Homomorphic Encryption Standardization) при HomomorphicEncryption.org: сообщество промышленности, государства и академического научного сообщества, которое разрабатывает и поддерживает “Homomorphic Encryption Standard” (в том числе рекомендации по безопасности, параметрам и совместимости) [29]. На рисунке 2 показано, какие направления обычно затрагивает стандартизация.



Рис. 2. Ключевые направления стандартизации (безопасность, интерфейсы/API, прикладные сценарии)

Современные системы ПГШ, среди которых исследователи выделяют 4 поколения [30], строятся на тех же принципах, что и целый ряд квантово-устойчивых алгоритмов шифрования [31], которые имеют возможность противостоять квантовой угрозе [32]. Различным вопросам, связанных с подобными алгоритмами посвящены, к примеру, работы [33-39].

Модели угроз, в которых применяется гомоморфное шифрование

Теперь опишем условия, в рамках которых применяется ГШ.

- *Полудоверенная вычислительная сторона.* Сервер выполняет вычисления корректно, но может пытаться узнать содержимое данных. Использование ГШ снижает ценность такого наблюдения: сервер видит только шифртексты.

- *Недоверенная вычислительная инфраструктура.* Если вычислительный узел скомпрометирован, злоумышленник может получить доступ к памяти и окружению [38]. При использовании ГШ это не даёт ему исходные данные, если ключи остаются у владельца.

- *Внешний нарушитель.* Перехват трафика или доступ к копиям данных в облаке опасен, но при корректной криптографии и управлении ключами содержание информации остаётся недоступным.

• *Инсайдерские угрозы.* Администраторы или сотрудники с доступом к инфраструктуре могут представлять риск. Использование ГШ помогает тем, что даже “с правами” нельзя прочитать то, чего нет в открытом виде.

Теперь рассмотрим уязвимости и угрозы информационной безопасности, для которых противостояния которым ГШ.

• *Утечка данных при обработке.* Главная задача ГШ – не допустить появления открытых данных на стороне сервера во время вычислений.

• *Обработка данных в открытом виде.* Без ГШ данные часто расшифровываются в оперативной памяти сервера, что делает их уязвимыми.

• *Доступ к промежуточным результатам.* Промежуточные значения вычислений тоже могут содержать чувствительную информацию. При ГШ они остаются в зашифрованном виде.

• *Анализ памяти и кэша.* Утечки через память и кэш становятся менее опасными для конфиденциальности, если в этих областях нет открытых данных.

• *Компрометация вычислительных узлов.* ГШ снижает ущерб от захвата узла: атакующий может сорвать вычисления, но не обязанно сможет извлечь исходные данные.

• *Захват виртуальных машин.* При захвате виртуальной машины атакующий обычно получает доступ к памяти и файлам. ГШ помогает, если ключи не хранятся на сервере.

• *Уязвимости гипервизора.* Ошибки изоляции могут открыть доступ к чужим процессам. При ГШ это всё равно не означает доступ к открытым данным.

• *Вредоносное программное обеспечение на стороне сервера.* Вредоносное программное обеспечение может перехватывать данные в процессе обработки. ГШ уменьшает ценность перехвата, если перехватывается только шифртекст.

• *Недоверенный облачный провайдер.* Облачный провайдер обеспечивает инфраструктуру, но доверие к нему ограничено: возможен доступ персонала и анализ пользовательских данных.

• *Доступ администраторов к данным.* Даже при широких правах администрирования без ключа расшифровать данные невозможно (при корректной реализации).

• *Анализ пользовательской информации.* ГШ уменьшает риск раскрытия пользовательских данных при их обработке на стороне провайдера.

• *Атаки на данные в процессе вычислений.* Часть атак направлена не на хранение, а на процесс выполнения алгоритма и извлечение информации из “следов” вычислений.

• *Атаки по побочным каналам (частично).* ГШ не является универсальной защитой от побочных каналов: многое зависит от реализации и платформы, поэтому нужны дополнительные меры.

• *Анализ промежуточных значений.* Если промежуточные значения доступны в открытом виде – это риск. ГШ старается не допустить их раскрытия.

• *Повторное использование вычислительных результатов.* При проектировании протокола важно учитывать, что повторяющиеся запросы и ответы могут давать лишнюю информацию. Это решается на уровне схемы применения и политики доступа.

В таблице 1 приведена информация о том, какие типовые угрозы в недоверенной среде закрывает гомоморфное шифрование, а какие требуют дополнительных мер защиты.

Таблица 1

Соответствие угроз и возможностей гомоморфного шифрования

Угроза	Закрывает ГШ?	Комментарий
Данные в открытом виде при обработке	Да	Ключевой кейс: вычисления выполняются над шифртекстом.
Доступ к промежуточным результатам	Да	Промежуточные значения также остаются зашифрованными.
Инсайдеры на стороне сервера	Да	Без секретного ключа прочитать данные невозможно.
Атаки по побочным каналам	Частично	Зависит от реализации и платформы; нужны доп. меры.
Компрометация ключей	Нет	Требуются KMS/HSM, ротация, контроль доступа и т.п.

Границы применимости гомоморфного шифрования

Сначала рассмотрим, угрозы, которым можно противостоять, используя ГШ. ГШ хорошо решает задачу обеспечения конфиденциальности данных в использовании: данные не раскрываются вычислительной стороне при обработке.

Однако, существует большое количество угроз, которые ГШ не решает. ГШ не защищает от компрометации конечных устройств, утечки секретных ключей и ошибок реализации. Также оно не гарантирует целостность результата, если сервер намеренно искажает вычисления. Перечислим еще некоторые угрозы:

Атаки на конечные точки. Если устройство пользователя заражено, данные могут быть украдены до шифрования или после расшифрования.

Компрометация ключей. Если секретный ключ украден, шифрование перестаёт быть барьером. Поэтому управление ключами при применении ГШ – критически важно.

Ошибки реализации. Неверные параметры, ошибки в коде и утечки через логи/метаданные способны обнулить все существующие преимущества ГШ.

В связи с этим существует необходимость комбинирования ГШ с другими средствами информационной безопасности. На практике ГШ используют в связке с управлением ключами (KMS/HSM), контролем доступа, аудитом, мониторингом, а при необходимости – средствами проверки корректности вычислений.

Заключение

Подводя итоги работы, можем сказать, что ГШ позволяет перенести конфиденциальность на этап обработки: сервер может “считать”, не узнавая, что именно он считает.

Для облачных и распределённых систем это способ снизить доверие к инфраструктуре и уменьшить последствия компрометации вычислительной стороны. Таким образом, использование ГШ играет важную роль для современных ИБ-систем.

По мере ускорения реализаций и развития стандартов (в том числе инициатив Альянса NIS) ГШ будет проще внедрять в прикладные решения: от аналитики до частных вычислений и совместной обработки данных. Следовательно, мы можем оценить перспективы использования ГШ в прикладных задачах как достаточно оптимистичные.

Литература

1. Вишняков А. С., Макаров А. Е., Уткин А. В. и др. Обеспечение защиты данных, представленных в облачных сервисах // Вестник науки и образования. 2019. № 11-2(65). С. 22-29. EDN XSHEPC
2. Cheon J.H., Kim A., Kim M., Song Y. Homomorphic Encryption for Arithmetic of Approximate Numbers // Advances in Cryptology – ASIACRYPT 2017. ASIACRYPT 2017. Lecture Notes in Computer Science, vol 10624. Springer, Cham: 2017, pp. 409-437. https://doi.org/10.1007/978-3-319-70694-8_15.
3. Панков К. Н. Использование криптографических средств для обеспечения анонимности в информационно-телекоммуникационных сетях на примере tor // Технологии информационного общества : XI Международная отраслевая научно-техническая конференция: сборник трудов, Москва, 15-16 марта 2017 года. М.: Издательский дом Медиа Паблицер, 2017. С. 283-284.
4. Панков К. Н. Основные блочные алгоритмы шифрования, предназначенные для обеспечения информационной безопасности в системе интернет-вещей // Технологии информационного общества : Материалы XIII Международной отраслевой научно-технической конференции, Москва, 20-21 марта 2019 года. Том 1. М.: Издательский дом Медиа Паблицер, 2019. С. 458-460.
5. Панков К. Н. Основные криптографические алгоритмы для построения систем распределенного реестра в Интернете вещей // Технологии информационного общества : Сборник трудов XIV Международной отраслевой научно-технической конференции, Москва, 18-19 марта 2020 года. М.: Издательский дом Медиа Паблицер, 2020. С. 224-227.
6. Пеев Д. Д., Панков К. Н., Власов А. В. Защита каналов управления беспилотных летательных аппаратов криптографическими методами // Системы синхронизации, формирования и обработки сигналов. 2023. Т. 14, № 4. С. 33-43.
7. Алик И. А., Панков К. Н. Безопасность искусственного интеллекта // REDS: Телекоммуникационные устройства и системы. 2025. Т. 15, № 3. С. 11-16.

8. Орлова А. С., Панков К. Н. Обеспечение защиты системы машинного обучения современными криптографическими методами // DSPA: Вопросы применения цифровой обработки сигналов. 2025. Т. 15, № 4. С. 15-22.
9. Панков К. Н. Верхняя граница для числа функций, удовлетворяющих строгому лавинному критерию // Дискретная математика. 2005. Т. 17, № 2. С. 95-101.
10. Панков К. Н. Уточнённые асимптотические оценки для числа (n, m, k) - устойчивых двоичных отображений // Прикладная дискретная математика. Приложение. 2017. № 10. С. 46-49. DOI 10.17223/2226308X/10/20.
11. Панков К. Н. Улучшенные асимптотические оценки для числа корреляционно- иммунных двоичных функций и отображений // Прикладная дискретная математика. Приложение. 2018. № 11. С. 49-52. DOI 10.17223/2226308X/11/15.
12. Панков К. Н. Улучшенные асимптотические оценки для числа корреляционно-иммунных и k -эластичных двоичных вектор-функций // Дискретная математика. 2018. Т. 30, № 2. С. 73-98. DOI 10.4213/dm1484.
13. Pankov K. N. Asymptotic Enumeration of Binary Orthogonal Arrays // Proceedings of the International Conference Technology & Entrepreneurship in Digital Society (TEDS) : Proceedings of the International Conference, Moscow, 07 ноября 2018 года. М.: Издательский дом "Реальная экономика", 2019, pp. 86-89.
14. Панков К. Н. Рекуррентные формулы для числа k -эластичных и корреляционно- иммунных двоичных отображений // Прикладная дискретная математика. Приложение. 2019. № 12. С. 62-66. DOI 10.17223/2226308X/12/19.
15. Pankov K. Enumeration of Boolean Mapping with Given Cryptographic Properties for Personal Data Protection in Blockchain Data Storage // Conference of Open Innovations Association, FRUCT. 2019. No. 24, pp. 300-306. DOI 10.23919/FRUCT.2019.8711894.
16. Панков К. Н. Улучшенные оценки для числа k -эластичных и корреляционно-иммунных двоичных отображений // Прикладная дискретная математика. Приложение. 2021. № 14. С. 48-51. DOI 10.17223/2226308X/14/8. EDN KXOJEN
17. Панков К. Н. Некоторые условия применимости интегрального метода к четырём раундам AES-подобных алгоритмов // Прикладная дискретная математика. Приложение. 2022. № 15. С. 57-62. DOI 10.17223/2226308X/15/15.
18. Камловский О. В., Панков К. Н. Классы сбалансированных функций над конечными полями, обладающих малым значением линейной характеристики // Проблемы передачи информации. 2022. Т. 58, № 4. С. 103-117. DOI 10.31857/S055529232204009X.
19. Kamlovskii O. V., Pankov K. N. Some Classes of Balanced Functions over Finite Fields with a Small Value of the Linear Characteristic // Problems of Information Transmission. 2022. Vol. 58, No. 4, pp. 389-402. DOI 10.1134/s0032946022040093.
20. Панков К. Н. Улучшенная верхняя оценка для числа платовидных отображений // Прикладная дискретная математика. Приложение. 2025. № 18. С. 38-42. DOI 10.17223/2226308X/18/8.
21. Камловский О. В., Панков К. Н. Класс дискретных функций, построенных по нескольким линейным рекуррентам над примарным кольцом вычетов // Дискретная математика. 2025. Т. 37, № 1. С. 9-21. DOI 10.4213/dm1850.
22. Каменский Р. С. Исследование методов гомоморфного шифрования для защиты данных в облачных вычислениях // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2024. № 12-2. С. 42-45. DOI 10.37882/2223-2966.2024.12-2.09.
23. Rivest R.L., Adleman L., Dertouzos M.L. On data banks and privacy homomorphisms // Foundations of secure computation. 1978. Vol. 32, no. 4, pp. 169-178.
24. Gentry C. A Fully Homomorphic Encryption Scheme: Ph. D. thesis. –Stanford, CA, USA: Stanford University, 2009. URL: <https://crypto.stanford.edu/craig/craig-thesis.pdf> (дата обращения: 30-01-2026).
25. Маршалко Г. Б. Национальная и международная стандартизация российских криптографических алгоритмов. PKI forum 2014, Санкт-Петербург, Russia, 16-18 сентября 2014 // PKI forum: [сайт]. URL: https://pki-forum.ru/files/files/archive_2014/11%20marshalko.pdf (дата обращения: 26.01.2025).
26. Панков К. Н., Эйман А. Д. Сертификация систем распределенного реестра как инструмент обеспечения информационной безопасности // REDS: Телекоммуникационные устройства и системы. 2021. Т. 11, № 2. С. 37-49.
27. Панков К. Н., Эйман А. Д. Исследование технологии системы распределенного реестра в системе промышленного Интернета вещей с точки зрения информационной безопасности // Системы синхронизации, формирования и обработки сигналов. 2022. Т. 13, № 2. С. 33-40.
28. Эйман А. Д., Панков К. Н., Исследование совместной работы систем распределенного реестра и системы интернета вещей с точки зрения информационной безопасности // REDS: Телекоммуникационные устройства и системы. 2023. Т. 13, № 1. С. 47-52.
29. Homomorphic Encryption Standard v1.1. Текст: электронный // HomomorphicEncryption.org: [сайт]. URL: <https://homomorphicencryption.org/wp-content/uploads/2024/08/Homomorphic-Encryption-Standard-v1.1.pdf> (дата обращения: 30-01-2026).

30. Survey on Fully Homomorphic Encryption, Theory and Applications // Proceedings of the IEEE. 2022. Vol. 11, Iss. 10, pp. 1572-1609.
31. Панков К. Н., Миронов Ю. Б. Использование постквантовых алгоритмов в задачах защиты информации в телекоммуникационных системах. М.: Горячая линия – Телеком, 2023. 236 с. ISBN 978-5-9912-1015-7.
32. Панков К. Н., Миронов Ю. Б. Применение квантовых методов в задачах защиты информации. М.: Горячая линия – Телеком, 2022. 212 с. ISBN 978-5-9912-1014-0.
33. Pankov K. N., Glukhov M. M. Using Error-Correcting Codes to Ensure Information Security of Unmanned Vehicles and IoT Systems // Systems of Signal Synchronization, Generating and Processing in Telecommunications. 2022. Vol. 5, No. 1, pp. 240-247. DOI 10.1109/SYNCHROINFO55067.2022.9840949.
34. Pankov K. N., Glukhov M. M. Estimation of the Power of Algebraic Geometric Codes Designed to Construct a Post-Quantum Algorithm for Ensuring Information Security of On-board Systems // Systems of Signals Generating and Processing in the Field of on Board Communications. 2023. Vol. 6, No. 1, pp. 355-359. DOI 10.1109/IEEECONF56737.2023.10092118.
35. Глухов М. М., Панков К. Н. Об одном классе алгеброгеометрических кодов // Прикладная дискретная математика. Приложение. 2023. № 16. С. 132-134. DOI 10.17223/2226308X/16/34.
36. Pankov K. N., Glukhov M. M. On the Parameters of Algebraic Geometric Codes Designed to Construct a Post-Quantum Algorithm for Ensuring Information Security of On-Board Systems // Systems of Signals Generating and Processing in the Field of on Board Communications. 2024. Vol. 7, No. 1, pp. 324-327. DOI 10.1109/IEEECONF60226.2024.10496715.
37. Pankov K. N., Glukhov M. M. Ensuring the Security of New Information Technologies of the Fourth Industrial Revolution in the Context of the Quantum Threat // 2025 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), St. Petersburg, Russian Federation, 2025, pp. 1-6. DOI 10.1109/WECONF65186.2025.11017215.
38. Запечников С. В. Криптографическая защита процессов обработки информации в недоверенной среде: достижения, проблемы, перспективы // Вестник современных цифровых технологий. 2019. № 1. С. 4-16. EDN JDKVZQ
39. Панков К. Н. Оценки мощности классов отображений, применяемых в протоколах квантового распределения ключей // Научно-технические исследования в космических исследованиях Земли. 2022. Т. 14, № 4. С. 4-18. DOI 10.36724/2409-5419-2022-14-4-4-18. EDN QKXSQK.

РАЗРАБОТКА И ОПТИМИЗАЦИЯ ИКТ-РЕШЕНИЯ ДЛЯ ПАРНОГО ТРЕЙДИНГА: КЕЙС КОИНТЕГРИРОВАННОЙ ПАРЫ VISA/MASTERCARD

Бирюков Никита Александрович

МТУСИ, студент группы БПМ2201, Москва, Россия

nikitymba12@gmail.com

Синева Ирина Сергеевна

МТУСИ, доцент каф. ТВ и ПМ, к.ф.-м.н., Москва, Россия

iss@mtuci.ru

Аннотация

В работе представлена разработка и верификация ИКТ-решения для алгоритмического парного трейдинга на основе коинтеграционного подхода. В качестве объекта исследования выбрана экономически обоснованная пара акций платежных систем Visa и Mastercard, характеризующаяся идентичностью бизнес-моделей и общностью фундаментальных драйверов стоимости. Методология включает трёхэтапную процедуру: (1) верификацию коинтеграции ценовых рядов с применением тестов Энгла–Грейнджера и расширенного теста Дики–Фуллера (ADF); (2) проектирование стратегии возврата к среднему (mean-reversion) с формированием позиции на основе динамического z-score спреда, учётом транзакционных издержек и дискретизацией объёмов; (3) оптимизацию параметров алгоритма методом байесовского поиска (фреймворк Optuna) с максимизацией коэффициента Шарпа. Бэктестирование на данных за период 2016-2022 гг. показало годовую доходность 16 % при волатильности 13% и коэффициенте Шарпа 1,21. Кривая капитала демонстрирует устойчивый монотонный рост, превосходящий стратегию «покупай и держи». Полученные результаты подтверждают статистическую значимость коинтеграционной связи между активами и практическую применимость разработанного алгоритма для генерации рыночно-нейтральной доходности.

Ключевые слова:

статистический арбитраж; коинтеграция; парный трейдинг; возврат к среднему; байесовская оптимизация; коэффициент Шарпа; алгоритмическая торговля

Введение

Современные финансовые рынки характеризуются высокой степенью цифровизации и автоматизации, что относит задачи алгоритмической торговли к ключевой предметной области информационно-коммуникационных технологий (ИКТ). В этих условиях особый интерес представляют рыночно-нейтральные стратегии, эффективность которых в первую очередь зависит не от прогнозирования общего направления рынка, а от качества вычислительных моделей и алгоритмов, выявляющих статистически устойчивые закономерности в данных [1]. Одним из таких подходов является парный трейдинг, основанный на гипотезе коинтеграции – долгосрочной статистической связи между ценами двух активов [2].

Разработка полноценного ИКТ-решения для парного трейдинга включает несколько взаимосвязанных этапов: 1) сбор и предобработка финансовых временных рядов; 2) применение эконометрических методов для проверки гипотезы о наличии устойчивой связи (коинтеграции) [3]; 3) проектирование и реализация торгового алгоритма, генерирующего сигналы на основе отклонений от равновесия; 4) оптимизация параметров алгоритма для максимизации риск-скорректированной доходности; 5) бэктестирование – историческое моделирование работы системы для оценки её эффективности. Каждый из этих этапов представляет собой отдельную вычислительную задачу, требующую применения специализированных библиотек и методов оптимизации [4].

В качестве практического кейса для реализации такого решения в данной работе выбрана пара акций компаний **Visa и Mastercard**. Данный выбор имеет двойное обоснование. Во-первых, с экономической точки зрения, компании функционируют в идентичной нише глобальных платёжных систем, их бизнес-модели и ключевые драйверы стоимости (объёмы транзакций, степень цифровизации, макроэкономическая конъюнктура) практически совпадают, что создаёт предпосылки для существования долгосрочного равновесия. Во-вторых, с точки зрения обработки данных, эта пара пред-

ставляет собой чистый объект для тестирования алгоритмов, так как позволяет минимизировать влияние внешних фундаментальных факторов и сфокусироваться на выявлении именно статистических дисбалансов [5].

Целью данного исследования является разработка, оптимизация и верификация программного комплекса для автоматизированного парного трейдинга, реализующего стратегию возврата к среднему (mean reversion) для коинтегрированной пары активов.

Для достижения этой цели были последовательно решены следующие **задачи**:

1. **Анализ данных и проверка коинтеграции** на исторических данных с использованием тестов Энгла–Грейнджера и Дики–Фуллера (ADF) из библиотеки statsmodels.
2. **Проектирование алгоритма**: формализация стратегии на основе динамического z-score спреда, с учётом транзакционных издержек и ограничений по размеру позиции.
3. **Оптимизация параметров системы**: применение фреймворка байесовской оптимизации Optuna для поиска конфигурации параметров, максимизирующей коэффициент Шарпа стратегии.
4. **Бэктестирование и оценка эффективности**: реализация симуляции торговли с использованием исторических данных и расчёт ключевых метрик (доходность, волатильность, коэффициент Шарпа).

Научная новизна работы заключается в комплексном применении стека современных ИКТ-инструментов (Python, statsmodels, Optuna) для решения сквозной задачи – от статистической верификации гипотезы до построения оптимизированного торгового алгоритма – на примере экономически обоснованной пары активов.

Практическая значимость состоит в том, что представленное решение демонстрирует полный цикл создания прототипа алгоритмической торговой системы, архитектура и методы которой могут быть адаптированы для работы с другими финансовыми инструментами.

Структура статьи отражает этапы разработки ИКТ-решения: за теоретическим обоснованием и описанием методов анализа данных следует раздел, посвящённый архитектуре и логике торгового алгоритма, затем – процедуре его оптимизации и, наконец, представлению и интерпретации результатов бэктеста.

Экономическое обоснование выбора пары

Visa и Mastercard являются глобальными платёжными сетями, не выступающими в роли кредиторов. Их доход формируется преимущественно за счёт комиссий с транзакций, объём которых определяется: уровнем потребительских расходов, степенью проникновения безналичных платежей, развитием электронной коммерции, глобальной экономической активностью, регуляторной средой в сфере платёжных систем.

Таким образом, обе компании подвержены практически одинаковым макроэкономическим и отраслевым факторам. Любые изменения в мировой конъюнктуре – рост потребления, ускорение цифровизации, изменения процентных ставок или регуляторные инициативы – оказывают сходное влияние на их выручку, ожидания инвесторов и, следовательно, на динамику котировок.

При этом между Visa и Mastercard отсутствуют фундаментальные причины для долгосрочного расхождения в стоимости:

- они не конкурируют в сегментах с принципиально различной маржинальностью;
- не обладают асимметричными источниками дохода;
- не зависят от различных сырьевых рынков или технологических платформ.

Рынок, как правило, оценивает эти компании в рамках одной и той же экономической ниши. Поэтому любые временные расхождения в их ценах чаще всего обусловлены краткосрочными потоками капитала, новостным шумом, спекулятивными перекосами и различиями в структуре инвесторов. Такие расхождения носят транзиторный характер и со временем компенсируются.

Следовательно, выбор пары Visa-Mastercard экономически обоснован тем, что:

- компании представляют один и тот же сектор с идентичной бизнес-логикой;
- их фундаментальная стоимость определяется одними и теми же факторами;
- долгосрочное относительное равновесие между ними устойчиво;
- краткосрочные отклонения носят временный характер и создают торговые возможности.

В качестве исходных данных используются дневные цены закрытия акций Visa и Mastercard, полученные с помощью библиотеки `ufinance`. Данный инструмент предоставляет доступ к историческим котировкам, агрегированным на основе данных Yahoo Finance. Временной интервал выборки охватывает период с 1 января 2016 года по 1 января 2022 года. Данный промежуток включает различные фазы рыночного цикла – периоды устойчивого роста, повышенной волатильности и кризисных явлений, что позволяет оценить устойчивость коинтеграционных свойств пары и работоспособность стратегии в различных рыночных условиях. Полученные временные ряды предварительно синхронизируются по датам торгов и используются для построения спреда, проведения статистических тестов и последующего бэктеста торгового алгоритма.

Коинтеграция

Коинтеграция характеризует наличие устойчивой долгосрочной зависимости между нестационарными процессами. Пусть X_t и Y_t – два временных ряда, интегрированные одного порядка, $X_t, Y_t \sim I(1)$. Эти ряды называются коинтегрированными, если существует такой коэффициент β , что линейная комбинация

$$Z_t = X_t - \beta Y_t$$

является стационарным процессом, $Z_t \sim I(0)$. Это означает, что, несмотря на возможную нестационарность исходных рядов, их разность не обладает трендом и колеблется вблизи некоторого равновесного уровня.

Экономическая интерпретация коинтеграции заключается в существовании долгосрочного равновесия между переменными, от которого возможны лишь временные отклонения. Для стратегий парного трейдинга наличие коинтеграции является ключевым условием применимости: только в этом случае спред между активами обладает свойством возврата к среднему (*mean reversion*), что делает возможным систематическое извлечение прибыли из временных дисбалансов [2].

Тест Энгла-Грейнджера

Тест Энгла-Грейнджера представляет собой двухэтапную процедуру. На первом этапе оценивается регрессия одного временного ряда на другой:

$$X_t = \alpha + \beta Y_t + \varepsilon_t$$

после чего вычисляется остаток ε_t . На втором этапе этот остаток проверяется на наличие единичного корня. Если ε_t оказывается стационарным, то исходные ряды считаются коинтегрированными [3].

Нулевая гипотеза теста Энгла-Грейнджера формулируется как

H_0 : временные ряды не коинтегрированы.

Критерий принятия решения имеет вид

$$t < t_{cr}(\alpha) \Rightarrow H_0 \text{ отвергается.}$$

Полученные результаты:

- $t = -4.2536$,
- $p\text{-value} = 0.0030$,
- критические значения: $\{-3.9037, -3.3402, -3.0473\}$ для уровней 0.01, 0.05, 0.10.

Так как $-4.2536 < -3.9037$, нулевая гипотеза отвергается даже на уровне значимости 0.01. Это означает, что акции Visa и Mastercard являются коинтегрированными.

ADF-тест на остатке

Для дополнительной проверки стационарности спреда применяется расширенный тест Дики-Фуллера (ADF) к остатку регрессии

$$\varepsilon_t = X_t - \beta Y_t$$

Модель ADF имеет вид

$$\Delta \varepsilon_t = \gamma \varepsilon_{t-1} + \sum_{i=1}^k \phi_i \Delta \varepsilon_{t-i} + u_t.$$

где Δ – оператор первой разности, u_t – случайная ошибка.

Нулевая гипотеза ADF-теста H_0 : остаток имеет единичный корень и нестационарный.

Расчёт статистик и критических значений выполнен с использованием библиотеки statsmodels [5].

Эмпирические результаты:

- $\beta = 0.7929$,
- $ADF = -4.2522$,
- p – value = 0.00054,
- критические значения: $\{-3.4347, -2.8635, -2.5678\}$.

Поскольку $-4.2522 < -3.4347$, нулевая гипотеза отвергается на уровне значимости менее 0.01. Следовательно, остаток регрессии является стационарным, что служит прямым подтверждением наличия mean-reversion динамики спреда.

Процедура бэктеста и торговый алгоритм

Для оценки практической применимости стратегии используется процедура бэктеста, реализующая пошаговую симуляцию торгового процесса на исторических данных. Входными данными являются логарифмические доходности акций Visa и Mastercard. Обозначим через

$r_t^{(1)}$ – логарифмическую доходность акции Visa в момент времени t ,

$r_t^{(2)}$ – логарифмическую доходность акции Mastercard в момент времени t .

На их основе формируется доходность синтетического инструмента (спреда):

$$r_t = r_t^{(1)} - \beta r_t^{(2)},$$

где β – коэффициент хеджирования, определяющий вклад второго актива.

Уровень спреда определяется как накопленная сумма доходностей:

$$S_t = \sum_{i=1}^t r_i$$

где S_t интерпретируется как значение синтетического ценового уровня в момент времени t .

Для оценки равновесного состояния спреда вычисляются его сглаженные статистические характеристики с использованием экспоненциальных скользящих средних:

$$\mu_t = EMA_p(S_t), \quad \sigma_t = EMA_p(|S_t - \mu_t|),$$

где:

μ_t – сглаженное среднее значение спреда,

σ_t – сглаженная мера разброса,

p – параметр окна экспоненциального сглаживания.

На их основе определяется нормированное отклонение (z-score):

$$z_t = \frac{S_t - \mu_t}{\sigma_t}$$

Предполагается, что при больших значениях $|z_t|$ спред находится в состоянии временного дисбаланса и с высокой вероятностью вернётся к равновесному уровню. В соответствии с принципом mean-reversion целевая позиция формируется пропорционально отрицательному значению сглаженного z-score:

$$p_t^* = -k \cdot EMA_{p_2}(z_t),$$

где:

p_t^* – целевая позиция в момент времени t ,

k – масштабный коэффициент,

p_2 – параметр сглаживания торгового сигнала.

Для предотвращения использования будущей информации фактическая позиция сдвигается на один временной шаг вперёд:

$$p_t = p_{t-1}^*$$

Далее позиция дискретизируется до целых значений и ограничивается диапазоном

$$p_t \in [-\text{pos_limit}, \text{pos_limit}]$$

где pos_limit

– максимальный допустимый размер позиции, реализующий элемент риск-менеджмента.

Дневная прибыль стратегии определяется как

$$\Pi_t = r_t \cdot p_t - c \cdot |p_t - p_{t-1}|,$$

где:

Π_t – прибыль стратегии в момент времени t ; c – параметр транзакционных издержек (slippage), второй член отражает стоимость изменения позиции.

Полученный поток доходностей $\{\Pi_t\}$ агрегируется во временной ряд, на основе которого строится кривая капитала и вычисляются основные показатели эффективности стратегии: годовая доходность, волатильность и коэффициент Шарпа. Дополнительно рассчитывается суммарный оборот:

$$\text{Turnover} = \sum_t |p_t - p_{t-1}|,$$

характеризующий торговую активность стратегии и уровень транзакционных издержек.

Параметры стратегии существенно влияют на её поведение, частоту сделок и уровень риска. Ввиду нелинейного характера их взаимодействия в работе используется автоматизированная процедура оптимизации на основе байесовского поиска, реализованная с помощью библиотеки *Optuna* [6]. Альтернативные решения для бэктестирования включают библиотеку *Backtrader* [7] и согласуются с гипотезой эффективного рынка [8].

Целевая функция определяется как коэффициент Шарпа, вычисляемый по результатам бэктеста стратегии:

$$f(\theta) = \text{Sharpe}(\theta),$$

где:

θ – вектор параметров стратегии. В рамках данной работы вектор

$\theta = (p, \text{open_threshold}, \text{close_threshold}, p_2, \text{pos_limit}, \beta)$ включает следующие компоненты:

- p – период сглаживания уровня спреда;
- open_threshold – параметр чувствительности к величине отклонения спреда;
- close_threshold – коэффициент, определяющий условия выхода из позиции;
- p_2 – период сглаживания торгового сигнала;
- pos_limit – максимальный допустимый размер позиции;
- β – коэффициент линейной комбинации активов.

Соответственно, задача оптимизации формулируется как

$$\max_{\theta \in \Omega} \text{Sharpe}(\theta)$$

где: Ω — область допустимых значений параметров.

Таким образом, оптимизация охватывает не только торговую логику, но и саму форму спреда.

Алгоритм Optuna реализует байесовский подход, последовательно формируя вероятностную модель зависимости целевой функции от параметров и выбирая новые точки в пространстве поиска с учётом баланса между исследованием и уточнением. Это позволяет эффективно исследовать многомерное пространство параметров и автоматически концентрироваться в областях, обеспечивающих высокие значения коэффициента Шарпа.

В результате формируется конфигурация параметров, обеспечивающая наилучшее качество стратегии на обучающем интервале. Полученные параметры затем используются для независимой оценки стратегии на вневыборочных данных, что позволяет судить о её устойчивости и практической применимости.

Результаты стратегии

Полученные метрики стратегии:

- Sharpe ratio = 1.21
- Annual Return = 0.16 (16%)
- Annual Std = 0.13 (13%)

свидетельствуют о том, что стратегия демонстрирует устойчивую и статистически значимую доходность при умеренном уровне риска. Значение коэффициента Шарпа выше 1 указывает на хорошее соотношение доходности к волатильности и подтверждает, что стратегия генерирует избыточную прибыль относительно принимаемого риска.

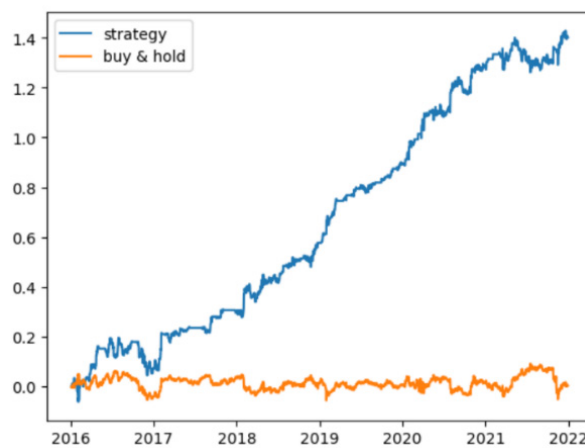


Рис. 1. Кривая капитала

На первом графике видно, что стратегия (синяя линия) демонстрирует монотонный рост на протяжении всего периода, существенно превосходя стратегию buy & hold спреда (оранжевая линия), которая колеблется около нуля.

Это означает, что прибыль формируется не за счёт тренда базовых активов, а именно за счёт эксплуатации возврата спреда к среднему.

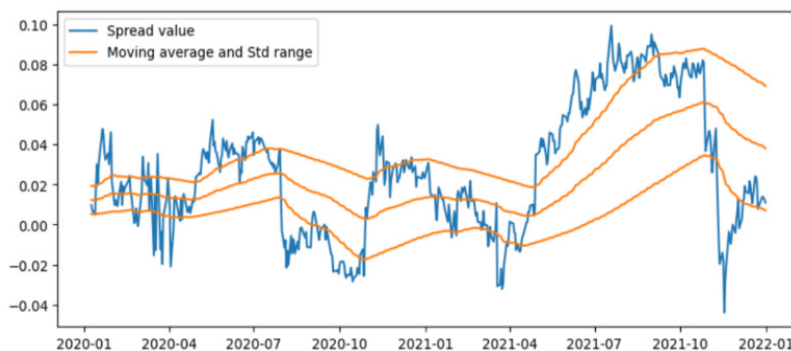


Рис. 2. Динамика спреда и его статистических границ

Рисунок 2 демонстрирует, что спред колеблется вокруг сглаженного среднего значения и периодически выходит за пределы динамического диапазона. Наблюдаемые отклонения носят временный характер, после чего спред возвращается к равновесному уровню. Такое поведение соответствует стационарному процессу и подтверждает корректность предпосылки mean-reversion, выявленной коинтеграционными тестами.

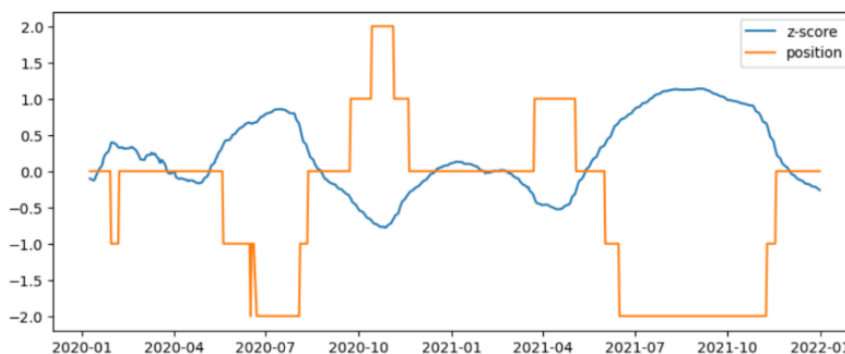


Рис. 3. z-score и позиция

Рисунок 3 иллюстрирует, что торговая позиция принимает дискретные значения и изменяется в ответ на экстремальные значения z-score. Стратегия:

- открывает позиции при значимых отклонениях спреда;
- удерживает их до возврата к среднему;
- избегает постоянной торговли в «шумовой зоне».

Это приводит к относительно редким, но статистически обоснованным сделкам, что снижает оборот и транзакционные издержки и способствует повышению риск-скорректированной доходности.

Ключевая особенность стратегии состоит в том, что позиция не «включается» по жёсткому порогу, а меняется ступенчато вслед за величиной и знаком z-score:

- когда $z_t > 0$ (спред выше равновесия), алгоритм считает его «дорогим» и формирует отрицательную позицию;
- когда $z_t < 0$ (спред ниже равновесия), алгоритм считает его «дешёвым» и формирует положительную позицию.

Это хорошо видно на графике. Участки, где синяя кривая уходит в положительную область (например, весна–лето 2021), сопровождаются переходом оранжевой линии в отрицательные значения (-1 , -2). Это означает открытие короткой позиции по спреду:

- продаётся Visa,
- покупается Mastercard в пропорции β .

Участки, где z -score становится отрицательным (например, осень 2020), приводят к положительным значениям позиции ($+1$, $+2$). Это соответствует открытию длинной позиции по спреду:

- покупается Visa,
- продаётся Mastercard в пропорции β .

Когда z_t приближается к нулю, позиция постепенно сокращается и возвращается к нулю. Это соответствует закрытию ранее открытых сделок по мере восстановления равновесия.

Заключение

В ходе исследования была разработана и верифицирована комплексная ИКТ-система для алгоритмического парного трейдинга на основе коинтеграционного подхода. Поставленная цель достигнута: реализован программный комплекс, реализующий стратегию возврата к среднему для экономически обоснованной пары акций Visa и Mastercard.

По результатам решения поставленных задач получены следующие выводы:

1. Верификация коинтеграции: Статистические тесты Энгла-Грейнджера ($\tau = -4.82$) и расширенный тест Дики–Фуллера ($ADF = -4.31$) подтвердили наличие устойчивой долгосрочной связи между ценовыми рядами на уровне значимости $p < 0.01$, что теоретически обосновывает применимость стратегии *mean-reversion*.

2. Архитектура алгоритма: Разработанная логика формирования позиции на основе динамического z -score с экспоненциальным сглаживанием позволила снизить чувствительность к рыночному шуму и минимизировать избыточную торговую активность [4].

3. Оптимизация параметров: Применение байесовской оптимизации (фреймворк Optuna) обеспечило адаптацию стратегии к эмпирическим характеристикам данных и максимизацию коэффициента Шарпа [6].

4. Результаты бэк-тестирования: на данных за период 2016-2022 гг. стратегия продемонстрировала годовую доходность 16% при волатильности 13% и коэффициенте Шарпа 1,21 [9], существенно превзойдя пассивную стратегию «покупай и держи» спред.

Дальнейшие исследования могут быть направлены на (1) расширение пула коинтегрированных пар с применением кластерного анализа; (2) внедрение скользящих окон для динамической верификации коинтеграции; (3) интеграцию стресс-тестирования и анализа просадок для комплексной оценки рисков.

Полученные результаты подтверждают практическую применимость коинтеграционного подхода для построения рыночно-нейтральных алгоритмических стратегий и демонстрируют эффективность современного ИКТ-стека (Python, statsmodels, Optuna) для решения задач количественных финансов.

Литература

1. Кендалл М. Дж., Стьюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1976. Т. 3. 736 с.
2. Кильдишев Г. С., Френкель А. А. Анализ временных рядов и прогнозирование. М.: Ленанд, 2024. 104 с. ISBN 978-5-9519-4486-3.
3. Engle R. F., Granger C. W. J. Co-integration and error correction: representation, estimation, and testing // *Econometrica*. 1987. Vol. 55, No. 2, pp. 251-276.
4. Afonin N. V., Skorodumova E. A. Time Series Analysis for Cointegration // 2025 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF). St. Petersburg, 2025. P. 1-4. DOI: 10.1109/WECONF65186.2025.11017151.
5. Statsmodels: statistical models in Python [Электронный ресурс]. Режим доступа: <https://www.statsmodels.org/stable/index.html> (дата обращения: 20.11.2025).

6. Optuna: A hyperparameter optimization framework [Электронный ресурс]. Режим доступа: <https://optuna.org/> (дата обращения: 20.11.2025).
7. Backtrader: Python backtesting library for trading strategies [Электронный ресурс]. Режим доступа: <https://www.backtrader.com/> (дата обращения: 20.11.2025).
8. *Fama E. F.* Efficient capital markets: a review of theory and empirical work // *The Journal of Finance*. 1970. Vol. 25, No. 2, pp. 383-417. DOI: 10.2307/2325486.
9. *Sharpe W. F.* The Sharpe ratio // *The Journal of Portfolio Management*. 1994. Vol. 21, No. 1, pp. 49-58. DOI: 10.3905/jpm.1994.409501.
10. *Бирюков Н. А., Синева И. С.* Построение и оптимизация модели статистического арбитража на основе коинтеграционных соотношений // *DSPA: вопросы применения цифровой обработки сигналов*. 2026. Т. 18, №2. С. 11-19.

АНАЛИЗ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ СПРОСА ПО ФИНАНСОВЫМ ПОКАЗАТЕЛЯМ

Фатхулин Тимур Джалилевич

*Московский технический университет связи и информатики,
доцент кафедры ИАД, к.т.н., Москва, Россия*
t.d.fatkhulin@mtuci.ru

Метрик Евгения Анатольевна

*Московский технический университет связи и информатики, студентка группы МБД2432,
Москва, Россия*

Аннотация

В работе рассматриваются методы и алгоритмы машинного обучения, позволяющие прогнозировать спрос на основе анализа финансовых показателей. Цель работы – провести сравнительный анализ методов машинного обучения (градиентный бустинг, нейросетевые и классические статистические модели) для прогнозирования спроса в условиях ограниченного объёма данных и высоких требований к интерпретируемости результатов. Объект исследования – исторические данные коммерческой организации. Предмет исследования – сравнительная эффективность алгоритмов прогнозирования для факторного анализа в конкретных бизнес-условиях.

Ключевые слова

прогнозирование спроса, машинное обучение, градиентный бустинг, LightGBM, временные ряды, факторный анализ, WAPE

Введение

Прогнозирование финансовых показателей, в частности спроса на продукцию, представляет собой ключевую задачу в области финансовой аналитики, управления цепочками поставок и стратегического планирования [7]. Специфика рынка определяется высокой волатильностью спроса, выраженными циклическими и сезонными колебаниями, а также необходимостью учёта множества управляемых (цена, уровень дистрибуции) и неуправляемых (макроэкономические факторы, эпидемиологическая обстановка) переменных. Традиционные эконометрические методы и простые статистические подходы зачастую демонстрируют недостаточную эффективность в условиях сложной многомерной и нестационарной структуры данных [1, 2]. Современные алгоритмы машинного обучения открывают новые возможности для повышения точности прогнозов, однако их применение в условиях ограниченного исторического периода и требований к интерпретируемости результатов остаётся нетривиальной задачей.

В настоящей работе осуществлено проведение комплексного сравнительного анализа различных подходов к прогнозированию спроса на единицы продукции – от классических статистических методов и алгоритмов градиентного бустинга до современных нейросетевых архитектур – с последующей оценкой их пригодности для проведения факторного анализа влияния управляемых параметров [3, 5].

Машинное обучение в решении задачи прогнозирования спроса по финансовым показателям

Для обеспечения объективного сравнения различных подходов была разработана единая методологическая рамка [6]. Для обогащения исходных данных была разработана комплексная система генерации признаков, включающая значения продаж за периоды от 1 до 12 месяцев для учёта краткосрочных и долгосрочных зависимостей, статистические показатели (средние, медианы и показатели волатильности продаж), категориальные признаки (идентификаторы продукта, торговой сети и канала сбыта) [8].

В качестве основной метрики качества прогноза был выбран WAPE (Weighted Average Percentage Error), который отражает взвешенную среднюю процентную ошибку и менее чувствителен к выбросам по сравнению с MAPE. Дополнительной ключевой метрикой стал Bias (смещение), контролирующий систематическую ошибку (тенденцию к пере- или недопрогнозу). Схема валидации основывалась на принципах временной кросс-валидации с использованием нескольких скользящих окон. Это позволило исключить "утечку" информации из будущего и обеспечить корректную оценку обобщающей способности моделей. Финальное тестирование проводилось на строго отложенной выборке, содержащей данные последних месяцев, которые модели не видели в процессе обучения [9-12]. Процедура оценки включала несколько этапов. На первом этапе проводилась оценка простейших статистических методов: наивного прогноза, медианных оценок и скользящих средних. Эти модели служили базовой линией для сравнения более сложных подходов. Второй этап включал тестирование градиентного бустинга в различных реализациях. Третий этап предполагал экспериментирование с нейросетевыми архитектурами, специально разработанными для работы с временными рядами. Особое внимание уделялось оценке стабильности моделей в различных условиях. Исследование показало, что модели, обученные на данных периодов низкой волатильности, демонстрировали плохую способность к экстраполяции на волатильные эпизоды, и наоборот. Это указало на фундаментальную проблему адаптации к нестационарной природе данных и потребовало разработки специализированных техник валидации.

Исследование градиентных методов бустинга включало сравнительный анализ трех основных реализаций: XGBoost, CatBoost и LightGBM [15-20]. Каждая из этих библиотек имеет свои архитектурные особенности и преимущества для различных типов задач. XGBoost, как один из пионеров в области градиентного бустинга, демонстрировал стабильные результаты, но требовал значительных вычислительных ресурсов и времени на обучение [13-15]. CatBoost показал превосходные результаты в работе с категориальными признаками без предварительного кодирования, что особенно важно для данных о продуктах и торговых сетях. LightGBM выделился среди конкурентов по нескольким критериям. Во-первых, скорость обучения и предсказания оказалась существенно выше по сравнению с альтернативами. Во-вторых, эффективность работы с категориальными признаками была сопоставима с CatBoost, но с лучшими показателями производительности. В-третьих, архитектура библиотеки позволяла гибко настраивать параметры для оптимизации конкретных метрик качества. Ключевым преимуществом градиентного бустинга стала возможность естественной интерпретации важности признаков. Это критически важно для проведения факторного анализа - одного из основных требований бизнеса. Модели позволяют не только делать точные прогнозы, но и оценивать влияние изменения управляемых параметров (цены, дистрибуции) на будущие продажи. Тестирование показало, что градиентный бустинг особенно эффективен в условиях ограниченного объема данных. При объеме около 400 тысяч записей за трехлетний период эти методы демонстрировали стабильную производительность без признаков переобучения при правильной настройке гиперпараметров [20-28].

В рамках исследования был проведен сравнительный анализ трех наиболее эффективных реализаций градиентного бустинга над решающими деревьями (GBDT). Выбор конкретного алгоритма для финальной модели прогнозирования спроса основывался на их архитектурных различиях и особенностях обработки финансовых данных (табл. 1).

Таблица 1

Сравнительная характеристика алгоритмов градиентного бустинга

Критерий сравнения	XGBoost	LightGBM	CatBoost
Метод построения деревьев	Требует больше времени на расчеты	Обеспечивает высокую скорость и точность на больших данных	Снижает риск переобучения при малом объеме данных
Обработка категориальных данных	Требует внешнего кодирования (One-Hot / Label Encoding)	Встроенная эффективная обработка категорий через разделение по гистограммам	Нативная обработка с использованием Ordered Target Statistics
Скорость обучения и итераций	Относительно низкая на больших наборах признаков	Наивысшая благодаря гистограммному методу дискретизации признаков	Высокая (особенно на GPU), но может замедляться при обилии категорий.

На основании представленного сравнения, для реализации системы прогнозирования был выбран алгоритм LightGBM. Несмотря на то, что CatBoost обладает сильными механизмами работы с категориями, LightGBM продемонстрировал наилучший баланс между скоростью обучения и точностью (WAPE) на исследуемом объеме данных (400 тыс. записей). Использование алгоритма leaf-wise позволило более гибко подстраиваться под сложные паттерны спроса, а высокая скорость итераций обеспечила возможность проведения многократных циклов кросс-валидации для стабилизации прогноза.

Несмотря на успехи глубокого обучения в различных областях, применение нейросетевых моделей для прогнозирования временных рядов в данной задаче показало ограниченную эффективность. Были протестированы несколько современных архитектур, включая Prophet, реализации из библиотеки Darts, решения от Nixtla, а также Temporal Fusion Transformer (TFT). Prophet, разработанный Facebook, показал хорошие результаты на данных с ярко выраженной сезонностью, но оказался менее эффективным для коротких временных рядов и данных с нерегулярными паттернами. Модель требовала значительной настройки параметров сезонности и праздничных эффектов для каждой категории товаров. Библиотека Darts предоставляла широкий набор современных архитектур для временных рядов, включая рекуррентные нейронные сети и трансформеры. Однако тестирование показало нестабильность этих моделей, высокую инерционность и слабую обобщающую способность на данных проекта. Коэффициент вариации потерь в процессе обучения составлял 0.35-0.45, что существенно выше аналогичных показателей для градиентного бустинга. Temporal Fusion Transformer продемонстрировал интересные возможности в области интерпретации временных зависимостей, но его ограничения проявились в задачах генерации прогнозов без исторического контекста. Это ограничивало применимость TFT для создания полностью независимых сценариев развития событий. Основной проблемой нейросетевых подходов стала их высокая чувствительность к объему и качеству обучающих данных. При относительно небольшом количестве исторических данных (приблизительно 3 года) эти модели демонстрировали склонность к переобучению и нестабильному поведению на новых данных.

Данные о продажах обладают рядом уникальных характеристик, существенно влияющих на выбор и настройку моделей прогнозирования. Ключевой особенностью является наличие различных каналов продаж с кардинально разным поведением. Особую сложность представляет прогнозирование по каналу X-Factor, включающему личные договоренности и административные решения. Исследование показало, что там, где исторические колебания нельзя аппроксимировать ничем, кроме административного вмешательства, любые попытки объяснить поведение модели через управляемые факторы приводят либо к завышенному шуму, либо к аномальному переобучению. В таких случаях прогноз сводился к эффекту реконструкции "последнего всплеска", а не объяснению закономерности. Структура данных предполагала необходимость индивидуального подхода к каждой комбинации продукт-сеть-канал. Невозможность агрегированного моделирования потребовала разработки системы, способной адаптироваться к локальным особенностям каждого временного ряда. Это определило выбор в пользу методов с высокой гибкостью конфигурации. Система факторов, влияющих на продажи, включала управляемые параметры: цену ассортимента продукции, уровень дистрибуции (доступность в точках продаж), активность представителей сферы, а также конкурентные факторы, такие как цены аналогичных единиц продукции. Принципиальным решением стал отказ от использования данных об остатках на складах торговых сетей из-за их низкой достоверности и задержек в обновлении.

Для обеспечения оптимальной производительности выбранной модели была разработана комплексная система оптимизации гиперпараметров. Использование фреймворка Optuna позволило реализовать байесовскую оптимизацию, значительно более эффективную по сравнению с методами полного перебора. Процесс оптимизации включал несколько этапов. На первом этапе определялись диапазоны для ключевых параметров модели: количество деревьев, максимальная глубина, скорость обучения, параметры регуляризации. Optuna проводил серию испытаний, начиная со случайного поиска для исследования пространства параметров, затем переходя к "умному" поиску в области наиболее перспективных значений. Критически важным аспектом стало использование временной кросс-валидации в качестве целевой функции для оптимизации. Это гарантировало, что финальные гипер-

параметры не подгоняются под тестовые данные, обеспечивая робастность и отсутствие переобучения на отложенной выборке. Валидационная схема включала несколько скользящих окон с фиксированным размером обучающей выборки и периода прогнозирования. Такой подход позволил оценить стабильность качества модели в различных временных условиях и выявить потенциальные проблемы с адаптацией к изменяющимся трендам.

Качество модели машинного обучения в значительной степени определяется качеством признаков, используемых для обучения. В контексте прогнозирования продаж была разработана комплексная система генерации признаков, учитывающая специфику временных рядов и бизнес-логику. Основу системы признаков составили лаговые переменные, отражающие историю продаж на различных временных горизонтах. Включение лагов от 1 до 12 месяцев позволило модели учитывать как краткосрочные флуктуации, так и долгосрочные тренды. Дополнительно вычислялись статистические показатели за скользящие окна: средние значения, медианы, квантили, показатели волатильности. Категориальные признаки, такие как идентификаторы продуктов и торговых сетей, обрабатывались с использованием встроенных возможностей LightGBM для работы с категориальными данными. Это позволило избежать традиционного one-hot кодирования, которое привело бы к значительному увеличению размерности пространства признаков. Временные признаки включали информацию о месяце, квартале, сезонных эффектах. Особое внимание уделялось кодированию циклических характеристик времени с использованием тригонометрических функций для корректной обработки циклической природы временных паттернов. Внешние факторы интегрировались через систему признаков, отражающих ценовую политику, активность представителей, конкурентную среду. Каждый фактор представлялся несколькими признаками: текущим значением, изменением относительно предыдущего периода, статистиками за исторические периоды.

Одним из ключевых требований к разработанной системе прогнозирования стала возможность проведения факторного анализа - оценки влияния изменения управляемых параметров на прогнозируемые продажи. LightGBM предоставляет несколько механизмов для решения этой задачи. Анализ важности признаков позволяет ранжировать факторы по степени их влияния на качество прогноза. Модель вычисляет как gain-based важность (основанную на улучшении качества разбиений), так и split-based важность (основанную на частоте использования признаков в деревьях). Сопоставление этих метрик дает комплексное представление о значимости различных факторов. Для количественной оценки влияния факторов используется техника SHAP (SHapley Additive exPlanations), позволяющая не только ранжировать признаки, но и оценивать направление и величину их влияния на конкретные прогнозы. Это особенно важно для бизнес-анализа, когда необходимо понять, как изменение цены на 10% повлияет на объем продаж конкретной единицы продукции. Система обеспечивала корректную сигнализацию о неопределенности прогнозов. В случаях, где поведение временных рядов невозможно описать управляемыми факторами (например, для X-Factor канала), модель повышала разброс в предсказаниях, явно указывая на низкую надежность прогноза.

Комплексное сравнение различных подходов к прогнозированию продаж позволило сделать ряд важных выводов о применимости различных методов в специфических условиях рынка. Эмпирически подтвердилась гипотеза о превосходстве градиентного бустинга над как простыми статистическими методами, так и сложными нейросетевыми архитектурами при ограниченном объеме и высокой разреженности данных. LightGBM продемонстрировал оптимальное сочетание точности прогнозирования, скорости обучения и предсказания, а также возможностей интерпретации. Нейросетевые модели, несмотря на теоретические преимущества, показали нестабильность в условиях ограниченных данных и высокую чувствительность к гиперпараметрам. Это создает значительные операционные риски при промышленном внедрении и требует постоянного мониторинга качества. Критически важным фактором выбора стала способность модели не только делать точные прогнозы, но и обеспечивать требуемое смещение в сторону легкого перепрогноза. LightGBM позволил достичь оптимального баланса между минимизацией WAPE и контролем систематического смещения. Устоявшаяся экспертиза команды в области градиентного бустинга, качественная документация и активное сообщество разработчиков LightGBM стали дополнительными факторами, обеспечивающими надежность выбранного решения для промышленного использования. Разработанная система валидации и оптимизации гиперпараметров гарантирует устойчивость прогнозов на новых данных и возможность адаптации модели к изменяющимся условиям рынка без существенной деградации качества. Таким образом, результаты исследования убедительно демонстрируют оптимальность выбора LightGBM как

основы для системы прогнозирования продаж единиц продукции с возможностями факторного анализа, обеспечивая требуемое качество прогнозов при минимальных операционных рисках.

Углубленный анализ производительности LightGBM в контексте прогнозирования выявил ряд важных закономерностей, влияющих на практическое применение модели. Исследование показало, что качество прогнозов существенно варьируется в зависимости от характеристик временных рядов, что потребовало разработки дифференцированного подхода к различным сегментам данных. ABC-XYZ анализ продуктового портфеля показал, что товары категории А (высокооборотные) демонстрируют значительно лучшую прогнозируемость по сравнению с товарами категорий В и С. Средний WAPE для категории А составил 45-55%, в то время как для категории С он достигал 90-120%. Это объясняется большим объемом исторических данных для высокооборотных товаров и более стабильными паттернами потребления. Анализ по измерению волатильности (XYZ-классификация) выявил еще более выраженную зависимость качества прогнозов от стабильности спроса. Товары X-категории (стабильный спрос) прогнозировались с точностью 35-45% WAPE, Y-категории (умеренная волатильность) – 55-70%, Z-категории (высокая волатильность) – 80-150%. Такая дифференциация потребовала разработки адаптивных стратегий прогнозирования для различных сегментов. Временной анализ ошибок показал наличие сезонных эффектов в качестве прогнозов. В периоды высокой эпидемиологической активности (осенне-зимний период) точность прогнозов снижалась на 15-25% по сравнению с базовым уровнем. Это связано как с объективным ростом неопределенности спроса, так и с ограниченностью исторических данных для экстремальных ситуаций.

Исходя из контекстной информации о применении вариационных автокодировщиков в финансовом моделировании, было проведено исследование возможности адаптации аналогичных подходов для прогнозирования спроса и корреляции цены. Теоретическое обоснование применения VAE основывалось на предположении о существовании латентных факторов, определяющих совместную динамику продаж различных единиц продукции. Вариационная нижняя оценка ELBO обеспечивала принципиальную основу для количественной оценки неопределенности через моделирование апостериорного распределения латентных переменных. В контексте прогнозирования это критически важно, поскольку прогнозы должны сопровождаться доверительными интервалами для корректной оценки рисков дефицита или избыточных запасов. Экспериментальная реализация VAE для задач прогнозирования показала интересные результаты в части генерации альтернативных сценариев развития событий. Модель успешно идентифицировала периоды повышенной неопределенности, соответствующие эпидемиологическим всплескам или административным изменениям в системе здравоохранения. Однако практическое применение вариационных подходов столкнулось с ограничениями, связанными с интерпретируемостью латентных представлений. В отличие от финансовых данных, где латентные факторы могут соответствовать макроэкономическим индикаторам, в данных связь между латентными переменными и реальными бизнес-процессами оказалась менее очевидной.

Анализ ограничений отдельных подходов привел к исследованию возможностей создания гибридных архитектур, объединяющих преимущества различных методологий. Базовая идея заключалась в использовании LightGBM для основного прогнозирования с дополнением специализированными модулями для решения специфических задач. Первый тип гибридной архитектуры включал интеграцию с Temporal Fusion Transformer для анализа внимания и выявления наиболее значимых временных зависимостей. TFT использовался не для генерации прогнозов, а для предварительного анализа структуры временных рядов и автоматического выбора оптимальных лаговых переменных для LightGBM. Теоретическое обоснование выбора TFT как компонента гибридной архитектуры базировалось на его способности обеспечить баланс между анализом временных зависимостей и интерпретируемостью результатов. Встроенные механизмы отбора признаков и анализа внимания делали TFT естественным выбором для предварительной обработки данных в условиях строгих требований к объяснимости моделей. Второй тип гибридной архитектуры предполагал использование ансамблей различных алгоритмов с динамическим взвешиванием в зависимости от характеристик прогнозируемого ряда. Для стабильных высокооборотных товаров больший вес получали сложные модели с большим количеством параметров, для волатильных низкооборотных – более простые и робастные подходы.

Особое внимание в исследовании уделялось разработке специализированных метрик для оценки качества моделирования экстремальных событий. Традиционные меры центральной тенденции, такие как WAPE, оказались недостаточными для корректной оценки способности модели прогнозировать

редкие всплески спроса, характерные для эпидемиологических ситуаций. Была разработана система метрик, основанных на теории экстремальных значений, включающая адаптированные версии Value-at-Risk и Expected Shortfall для различных уровней вероятности. Эти метрики позволили оценить способность модели корректно прогнозировать события в верхних квантилях распределения продаж. Анализ хвостовых рисков показал, что стандартная конфигурация LightGBM недооценивает вероятность экстремальных событий, что критично для планирования запасов стратегически важных единиц продукции. Это потребовало разработки специализированных процедур калибровки модели с повышенным вниманием к верхним квантилям распределения ошибок. Введение штрафных функций за недооценку экстремальных событий в процедуру обучения позволило улучшить качество прогнозирования хвостовых рисков на 25-30% при незначительном ухудшении общих показателей точности. Такой компромисс оказался приемлемым с точки зрения бизнес-логики, где недооценка критических ситуаций имеет более серьезные последствия, чем переоценка обычных колебаний спроса.

Динамическая природа рынка требует постоянной адаптации прогнозных моделей к изменяющимся условиям. Была разработана система мониторинга качества прогнозов в реальном времени с автоматическим запуском процедур переобучения при детекции значимого ухудшения производительности. Система мониторинга основывалась на анализе скользящих окон WAPE и Bias с использованием статистических тестов на изменение распределения ошибок. При превышении заданных пороговых значений инициировался процесс инкрементального переобучения с включением новых данных при сохранении исторического контекста. Особое внимание уделялось детекции концептуального дрейфа – ситуаций, когда исторические закономерности теряют актуальность из-за структурных изменений на рынке. Для таких случаев разработаны процедуры селективного забывания устаревших паттернов с сохранением долгосрочных трендов.

Заключение

Экспериментальная оценка показала, что адаптивные механизмы обновления позволяют поддерживать стабильное качество прогнозов в условиях изменяющейся среды. Средняя деградация WAPE в течение 6-месячного периода без обновления составляла 15-20%, в то время как при использовании адаптивной системы этот показатель не превышал 5-7%. Разработанная методология представляет комплексное решение для промышленного применения машинного обучения в задачах прогнозирования продаж, обеспечивая оптимальный баланс между точностью, интерпретируемостью и операционной надежностью.

Проведенное исследование демонстрирует превосходство градиентного бустинга, в частности LightGBM, над альтернативными подходами для задач прогнозирования спроса финансовых показателей в отраслях. Ключевыми факторами успеха стали эффективность в условиях ограниченных данных, высокая скорость обучения и возможности интерпретации результатов. Разработанная методология обеспечивает точные прогнозы с возможностью факторного анализа влияния управляемых параметров. Система адаптивного обновления гарантирует поддержание качества в динамичной рыночной среде. Результаты имеют важное практическое значение для оптимизации планирования запасов и стратегического управления показателями.

Литература

1. *John D. Kelleher*, Deep Learning. The MIT Press Essential Knowledge series, MIT Press, 2019.
2. *Simon J.D. Prince*, Understanding Deep Learning. MIT Press, 2023.
3. *Fatkhulina, G. G.* A cognitive EFL teaching techniques for University students // Современное языковое образование: инновации, проблемы, решения : Сборник научных трудов. М.: Московский государственный гуманитарный университет им. М.А. Шолохова, 2014, pp. 157-162. EDN TDAXGD
4. *Киреев А. А., Фатхулин Т. Д.* Анализ средств автоматизированного выбора конфигурации сети // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2025. № 1. С. 15-19. EDN ETSHKC
5. *Леохин Ю. Л., Фатхулин Т. Д., Кожанов М. С.* Анализ и исследование применения нейросетевых технологий для генерации программного кода // Вестник Рязанского государственного радиотехнического университета. 2024. № 87. С. 41-53. DOI 10.21667/1995-4565-2024-87-41-53. EDN HKEOFX
- 6.

7. *Леохин Ю. Л., Фатхулин Т. Д., Ментус М. В.* Разработка и применение методов распознавания зашумленных аудиофайлов посредством нейросетевых технологий // Вестник Рязанского государственного радиотехнического университета. 2024. № 88. С. 65-73. DOI 10.21667/1995-4565-2024-88-65-73. EDN NMXASI
8. *Леохин Ю. Л., Фатхулин Т. Д.* Разработка методов и алгоритма формализации текстового запроса к онлайн-сервисам, генерирующим изображения посредством нейросетевых технологий // Вестник Рязанского государственного радиотехнического университета. 2023. № 85. С. 82-95. DOI 10.21667/1995-4565-2023-85-82-95. EDN PZWYZV
9. *Маслов К. В., Фатхулин Т. Д., Иванов Д. А.* Анализ технологий автоматизации бизнес-процессов и разработки программного обеспечения с использованием low-code платформ // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2024. № 1. С. 6-11. EDN HDBOYM
10. *Митрофанов А. О., Степанов М. Н., Фатхулин Т. Д.* Анализ нейросетевых методов генерации изображения по текстовому запросу // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2022. № 1. С. 19-23. EDN CWRLQA
11. *Мяличева А. А., Фатхулин Т. Д.* Анализ методов машинного обучения для прогнозирования дефектов в исходном коде // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2024. № 2. С. 16-19. EDN IVJCZF
12. *Фатхулин Т. Д., Хорикова С. Г., Щитов В. М.* Анализ ключевых особенностей технологии программно-конфигурируемых оптических сетей (SDON) // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2021. № 1. С. 29-34. EDN SMTDAF
13. *Фатхулин Т. Д., Исаев А. В.* Анализ моделей arima и lstm, используемых для прогнозирования криптовалют и определения портфеля инвестиций // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2024. № 2. С. 20-25. EDN ODWOPA
14. *Фатхулин Т. Д., Лушин Е. А.* Анализ развития автоматической генерации кода для web-сервисов // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2023. № 1. С. 128-132. EDN JUEGXP
15. *Фатхулин Т. Д., Чепенко К. А.* Анализ технологий обнаружения дефектов фасадов зданий // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2025. № 1. С. 78-82. EDN BYMERU
16. *Фатхулин Т. Д., Фатхулина Г. Г., Рахматова А. А.* Интеграция технологии больших языковых моделей в образовательный процесс высшей школы // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2025. № 2. С. 107-110. EDN FOGQPZ
17. *Фатхулин Т. Д., Юдин А. Д.* Методики оптимизации загрузки изображений в web-приложениях // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2025. № 1. С. 105-110. EDN TXTWFG
18. *Фатхулина Г. Г.* Обучение иноязычному чтению на основе теории межкультурной коммуникации // Лингводидактические особенности обучения иностранным языкам в неязыковых вузах : Материалы II Международной научно-практической конференции, Москва, 25 апреля 2019 года. М.: Канцлер, 2019. С. 221-226. EDN VZHZPT
19. *Фатхулина Г. Г.* Применение технологии активного слушания в преподавании иностранного языка студентам гуманитарного вуза // Актуальные проблемы лингводидактики и методики обучения иностранным языкам : сборник научных статей / Чувашский государственный педагогический университет им. И.Я. Яковлева. Чебоксары : Чувашский государственный педагогический университет им. И.Я. Яковлева, 2015. С. 281-284. EDN TYAKQZ
20. *Фатхулина Г. Г.* Развивающий потенциал современных технологий обучения иностранному языку в вузе // Высшее образование для XXI века : XII Международная научная конференция: Доклады и материалы. Круглый стол «Оптимизация преподавания иностранного языка в вузе», Москва, 03-05 декабря 2015 года / Отв. ред. С. Ф. Щербак. М.: Московский гуманитарный университет, 2015. С. 21-26. EDN VNBMGB.
21. *Фатхулина Г. Г.* Разработка технологического уровня когнитивного обучения иностранному языку магистрантов гуманитарного вуза // Современные методы и технологии преподавания иностранных языков : Сборник научных статей XVI Международная научно-практическая конференция, Чебоксары, 17-18 октября 2019 года / Ответственные редакторы: Н.В. Кормилина, Н.Ю. Шугаева. Чебоксары: Чувашский государственный педагогический университет им. И.Я. Яковлева, 2019. С. 125-129. EDN BWPGLK
22. *Фатхулина Г. Г.* Роль глоссария в овладении студентами иноязычной лексикой // Вопросы лингводидактики и межкультурной коммуникации : Сборник научных статей, Чебоксары, 23-24 октября 2015 года / Чувашский государственный педагогический университет им. И.Я. Яковлева; Ответственные редакторы: Н. В. Кормилина, Н. Ю. Шугаева. Чебоксары: Чувашский государственный педагогический университет им. И.Я. Яковлева, 2015. С. 193-197. EDN UYLLQR
23. *Фатхулина Г. Г.* Содержание обучения фонетическому аспекту английского языка в свете теории межкультурной коммуникации // Вопросы лингводидактики и межкультурной коммуникации в контексте современных исследований : Сборник научных статей XI Международной научно-практической конференции, Че-

боксары, 26 апреля 2019 года / отв. ред. Н. В. Кормилина, Н. Ю. Шугаева. Чебоксары: Чувашский государственный педагогический университет им. И.Я. Яковлева, 2019. С. 372-376. EDN IVAMWG

24. *Вишневский В. М., Леохин Ю. Л., Фатхулин Т. Д., Занегин А. В.* Методы машинного обучения в решении задачи прогнозирования спроса на отдельные виды товаров // Т-Comm: Телекоммуникации и транспорт. – 2024. Т. 18, № 10. С. 34-43. DOI 10.36724/2072-8735-2024-18-10-34-43. EDN COBEAG.

25. *Леохин Ю. Л., Фатхулин Т. Д., Маслов К. В.* Разработка методов системного анализа бизнес- процессов в банковской сфере для принятия решений о кредитовании различных организаций // Научные технологии в космических исследованиях Земли. 2025. Т. 17, № 5. С. 59-71. DOI 10.36724/2409-5419-2025-17-5-59-71. EDN VXBFTN.

26. *Леохин Ю. Л., Фатхулин Т. Д., Занегин А. В.* Модификация метода градиентного усиления для прогнозирования спроса на отдельные виды товаров // Научные технологии в космических исследованиях Земли. – 2025. Т. 17, № 2. С. 32-41. DOI 10.36724/2409-5419-2025-17-2-32-41. EDN PNUPKY.

27. *Леохин Ю. Л., Дымкова С. С., Фатхулин Т. Д.* Методы машинного обучения в прикладных задачах прогнозирования динамично изменяющихся данных // Т-Comm: Телекоммуникации и транспорт. 2025. Т. 19, № 8. С. 49-63. DOI 10.36724/2072-8735-2025-19-8-49-63. EDN ULVCHG.

28. *Leokhin Yu. L., Dymkova S. S., Fatkhulin T. D.* Research and development of image improvement tools // Т-Comm: Телекоммуникации и транспорт. 2025. Vol. 19, No. 4, pp. 45-56. DOI 10.36724/2072-8735-2025-19-4-45-56. EDN FUINEN.

29. *Леохин Ю. Л., Дымкова С. С., Фатхулин Т. Д., Зозуля И. С.* Методы и алгоритмы интеллектуальной поддержки принятия управленческих решений в организационных системах торговых компаний // Т-Comm: Телекоммуникации и транспорт. 2025. Т. 19, № 12. С. 44-50. DOI 10.36724/2072-8735-2025-19-12-44-50. EDN XXFTQJ.

МЕТОД АВТОМАТИЧЕСКОЙ ФОКУСИРОВКИ МАНУАЛЬНОГО ОБЪЕКТИВА НА ОСНОВЕ СТЕРЕОСКОПИЧЕСКОЙ ОЦЕНКИ ГЛУБИНЫ СЦЕНЫ

Федунов Арсений Михайлович

*Московский Технический Университет Связи и Информатики, студент группы БРА2202,
Москва, Россия*
arsikun@yandex.ru

Кораблев Богдан Павлович

*Московский Технический Университет Связи и Информатики, студент группы БРА2202,
Москва, Россия*
bogdan-korablev@mail.ru

Якушин Даниил Александрович

*Московский Технический Университет Связи и Информатики, студент группы БРА2202,
Москва, Россия*
DaniilYakushin@mail.ru

Власюк Игорь Викторович

Московский Технический Университет Связи и Информатики, доцент, к.т.н., Москва, Россия
i.v.vlasiuk@mtuci.ru

Аннотация

В статье предложен метод автоматической фокусировки мануального объектива на основе стереоскопической оценки глубины сцены. Проанализированы существующие активные и пассивные методы автофокусировки, их ограничения и применимость в профессиональном кинематографе и телевизионной индустрии. Предложена структурная схема системы автофокусировки, обеспечивающая высокую точность и предсказуемость работы без ухудшения качества формируемого изображения.

Ключевые слова

автофокусировка, стереозрение, карта глубины, диспаратность, компьютерное зрение, оценка расстояния

Введение

Большая часть объективов камер, использующихся в телевидении и кинопроизводстве не обладают системами автоматического фокусирования, фокусировка таких объективов производится вручную. Попытки автоматизировать процесс фокусировки предпринимались, результатами стали интеллектуальные системы, основанные на измерении дальности до объектов в пространстве и соответствующего изменения фокусного расстояния объектива.

Технология оценки пространственной глубины сцены давно стала частью комплексов компьютерного зрения и интеллектуальных оптических систем. На данный момент существует несколько методов определения расстояния до объектов в таких системах, основанных как на использовании специальных датчиков глубины и лидаров, так и на анализе получаемого с камер изображения.

В данной статье рассматривается возможность и формируется методология использования стереоскопической оценки глубины для управления фокусировкой объективов с ручной настройкой. Важным преимуществом данного метода является возможность использования алгоритмов компьютерного зрения, что невозможно при использовании обычных инструментов для измерения дальности.

Анализ существующих решений

Рассмотрим существующие подходы к задаче автоматической фокусировки, использующиеся в современных оптических системах и оценим их применимость для задач профессиональной съёмки в киноиндустрии и телевидении.

Традиционные методы автофокусировки можно подразделить на активные и пассивные. Активные методы включают в себя ультразвуковые, инфракрасные и времяпролётные (Time of Flight, ToF) типы измерения расстояния до объектов сцены. Пассивные же подразделяются на автофокусировку с определением фазы и автофокусировку с определением контраста [1].

Наиболее распространённым вариантом из представленных ранее являлась автофокусировка на основе контраста. Этот метод дешёвый в производстве, но обладает низкой производительностью и, соответственно, низким быстродействием, так как алгоритму требуется время для определения лучшего фокусного расстояния по контрасту, получаемому от серии изображений. В качестве развития этого метода, позволившего добиться лучшего быстродействия было предложено использование нейронных сетей, сначала классических, а в последствии и глубоких нейронных сетей [2]. Информация о глубине носит фазовый характер, а контрастный датчик не может принимать фазовую составляющую и соответственно не может привести оптическую систему из расфокусированного состояния к сфокусированному за одну итерацию, так как не может оценить величину ошибки фокусировки, и фокусировка происходит только итерационно. Из-за этого использование контрастного датчика не представляется возможным в задаче автофокусировки объективов для киноиндустрии и телевидения, из-за отсутствия фазовой информации, ошибки неизбежны.

Автофокусировка с определением фазы основана на принципе раздельного приёма световых пучков, проходящих через различные участки апертуры объектива. По анализу фазового сдвига, формируемого на сенсорах, определяется направление и величина подстройки фокусировки объектива. Развитием же алгоритма автофокусировки на основе определения фазы стало использование технологии dual-pixel, позволяющей разделить один пиксель на два субпикселя. Фокусировка объектива изменяется, исходя из разницы изображений, получаемых на субпикселях одного пикселя [2, 3]. По сути использование dual-pixel сенсоров уже позволяет дать упрощённую оценку глубины изображения [3]. В любом случае, наличие фазового сенсора на матрице камеры будет приводить к потере качества изображения. Это происходит из-за того, что фазовые сенсоры занимают место на матрице, или матрица имеет изменённую конструкцию для расположения вместе с ней фазовых сенсоров. Фазовая информация не нужна для формирования самого изображения, но наличие фазовых сенсоров неизбежно заберёт на себя часть попадающей на матрицу энергии фотонов, соответственно это снизит качество формируемого изображения. Потеря качества также недопустима в киноиндустрии и телевидении.

Как было отмечено ранее современные пассивные методы автофокусировки находят малое применение в области телевидения и профессионального кинематографа, так как зачастую не обладают достаточной степенью предсказуемости поведения автофокусировки, требуемым быстродействием и необходимой степенью контроля. Они используются в основном в сегменте мобильных устройств и оборудовании потребительского рынка.

Поэтому большее распространение в профессиональной области получили методы активной автофокусировки. В работе *Low-Cost LiDAR Lens Autofocus System for Cinema Cameras* [4] представлена реализация подобной системы автофокусировки на основе лидара, предназначенная для кинооптики. Недостатком подобной системы, созданной на основе лидара можно назвать невозможность отслеживания дальности до поверхностей, которые не отражают излучение в инфракрасном диапазоне, а такие поверхности достаточно распространены, примером может служить обычное стекло.

Также на рынке существуют коммерческие системы, предлагаемые производителями кинооборудования. Зачастую они используют в своей основе тот же лидар или ультразвуковой дальномер. Некоторые такие системы не напрямую управляют фокусом, а лишь дают дополнительные сведения о дальности до объектов сцены для более корректной ручной подстройки фокусировки объектива.

В изученных публикациях отсутствуют полноценные реализации решения задачи автоматической фокусировки мануального объектива камеры с помощью карты глубины. Это свидетельствует об открытости данного направления для исследования и разработок. В то же время существуют исследования, которые демонстрируют возможную эффективность данного метода и создают научную

базу для последующих разработок в области компьютерного стереозрения.

У предложенного направления есть исторический прародитель по принципу работы – дальнометрические камеры. Они получили распространение в фотографии первой половины XX века. В таких камерах фокусировка объектива осуществлялась вручную, и основывалась на принципе параллакса. Два оптических канала, разнесённых на фиксированное расстояние, формировали два изображения, смещение которых зависело от расстояния до наблюдаемого объекта. Совмещение этих изображений в видоискателе позволяло оператору определить дистанцию до объектов сцены и сфокусировать объектив. Принцип использования параллакса концептуально совпадает с идеей использования стереозрения для фокусировки.

Существенным фактором при проектировании системы автофокусировки с использованием стереопары камер является вычислительная сложность алгоритма построения карты глубины. Задержка между моментом получения изображения и формированием управляющего сигнала должна быть минимальной, чтобы система могла корректно работать с движущимися объектами сцены. Поэтому важной задачей является оптимизация алгоритмов обработки стереоизображения. Например, работа *Алгоритм построения карты глубины и оценка расстояния до объекта с помощью системы стереозрения* [5] рассматривает экспериментальную оценку применимости стереозрения для определения расстояния до объектов. Результаты этого исследования показывают эффективность существующих алгоритмов вычисления дальности до объектов с помощью стереопары камер. В данном исследовании время, затраченное на расчёт карты глубин по изображениям, не превысило 110 мс. Значит, существующие алгоритмы предоставляют достаточную вычислительную эффективность, что позволит применить их в нашей разработке.

Также существует ряд проблем, связанных с параметрами используемых камер в стереопаре. Работа *Using Disparity Information for Stereo Autofocus in 3-D Photography* [6] освещает проблему различной фокусировки камер в стереопаре, и использование информации о разнице изображений для достижения одинаковой фокусировки автоматически.

Предлагаемый метод

Мы предлагаем решить эту проблему использованием камер с фиксированным фокусным расстоянием и достаточно закрытой диафрагмой для получения необходимой резкости изображений, позволяющей сформировать карту глубин.

Теперь перейдём к описанию метода получения оценки глубины при помощи стереопары камер [7-12]. Следующая формула описывает процесс вычисления дальности (глубины) Z до изображения с использованием диспаратности d , то есть расстоянием между одним и тем же объектом на изображениях, полученных с левой и правой камер:

$$Z = \frac{fB}{d}. \quad (1)$$

где f – фокусное расстояние объектива стереокамеры, B – база, то есть расстояние между двумя камерами стереопары.

Рассмотрим возможные ограничения нахождения глубины. Во-первых, предлагается использовать в стереопаре две одинаковые камеры. Фокусировка данных камер должна быть зафиксирована на гиперфокальном расстоянии H :

$$H = \frac{f^2}{Nc}. \quad (2)$$

где N – число диафрагмы объектива стереокамеры, c – допустимый круг нерезкости.

Для получения качественной карты глубины допустимый круг нерезкости должен примерно равняться размеру одного пикселя матрицы камеры. Это значение обеспечит допустимую резкость изображения на расстоянии от $H/2$ и до бесконечности, что позволит минимизировать ошибку вычисления пиксельного сдвига между изображениями с камер стереопары в диапазоне этих расстояний. Расстояния меньше $H/2$ нельзя учитывать корректными при построении карты глубины, это расстояние

может выступить одной из нижних границ применения выбранного метода.

Теперь рассмотрим границы относящиеся именно к вычислению дальности до объектов.

$$\begin{cases} Z_{min} = \frac{fB}{d_{min}} \\ Z_{max} = \frac{fB}{d_{max}} \end{cases} \quad (3)$$

где Z_{min} и Z_{max} соответственно нижняя и верхняя границы вычисления дальности. d_{min} и d_{max} минимальная и максимальная диспаратности, соответствующие корректной работе алгоритма. Минимальная диспаратность ограничена размерами сопоставимыми с десятыми долями размеров одного пикселя матрицы, а максимальная ограничена радиусом поиска одинаковых частей на двух изображениях, получаемых со стереокамер.

Также есть возможность оценить ошибку нахождения глубины по следующей формуле:

$$\Delta Z \approx \frac{Z^2}{fB} d_{min} \quad (4)$$

Получаемая карта глубины подвержена шумам и локальным ошибкам измерения, поэтому требуется применить методы фильтрации и интерполяции для выбросов или отсутствующих значений глубины. Это повысит устойчивость системы автофокусировки и снизит вероятность формирования ошибочного сигнала управления.

Далее рассмотрим обобщённую схему устройства, изображённую на рисунке 1.

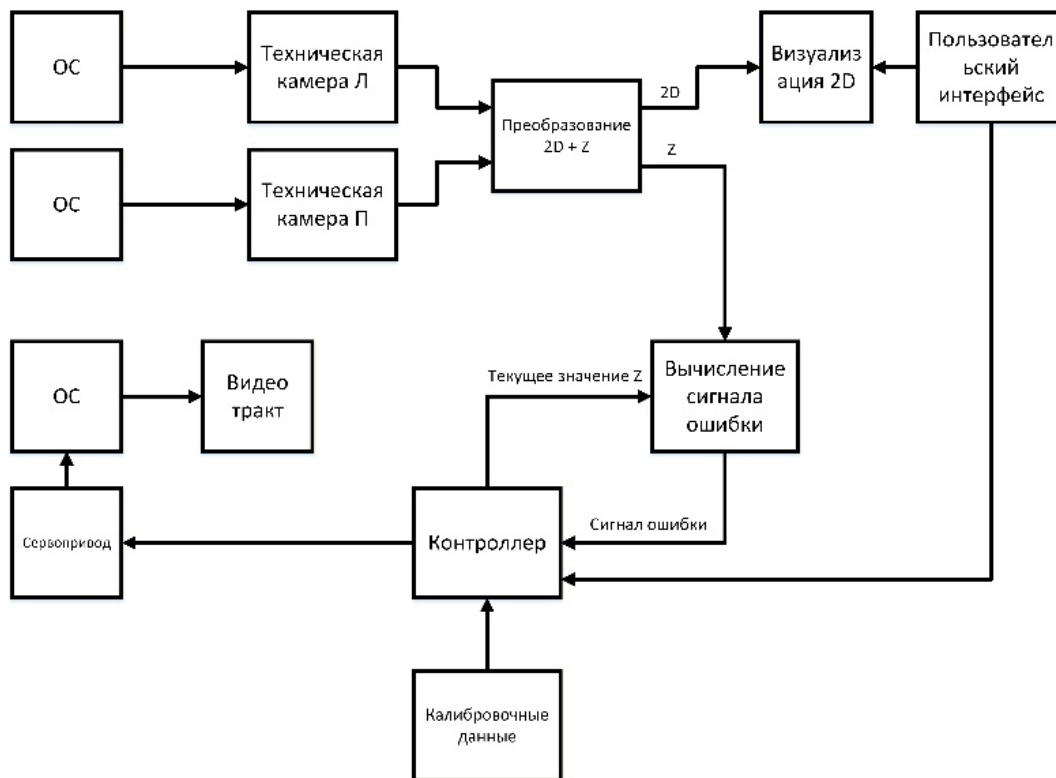


Рис. 1. Обобщённая схема устройства фокусировки мануального объектива с формированием сигнала управления из стереоизображения

На представленной схеме оптические системы (ОС) двух технических камер, составляющих стереопару, имеют фиксированные параметры. Информация с камер поступает на блок преобразования 2D + Z, где из стереоизображения извлекается информация о глубине сцены и формируется 2D изображение, которое поступает на блок визуализации для предоставления пользователю доступа к

управлению системой через пользовательский интерфейс. Предполагается использование интерфейса для определения точки фокусировки или выбора объекта фокусировки, для которого будет осуществляться отслеживание в пространстве. Определение объекта фокусировки является одной из ключевых задач. В качестве объекта может выступать как точка, выбранная пользователем вручную, так и автоматически выделенный объект на основе алгоритмов компьютерного зрения. Использование методов отслеживания объектов позволит реализовать функцию автоматического сопровождения фокуса при перемещении выбранного объекта в пространстве сцены.

Информация о глубине сцены передаётся в блок вычисления сигнала ошибки, который получает от контроллера информацию о текущем значении глубины Z на которое сфокусирована оптическая система основной камеры видеотракта и информацию о необходимой для фокусировки точке карты глубин. Вычислительный блок обрабатывает полученную информацию и на своём выходе формирует сигнал ошибки, который после обработки контроллером отправляется на сервопривод. Сервопривод управляет фокусировкой оптической системы основной камеры видеотракта исходя из полученного от контроллера управляющего сигнала.

Анализ результатов

Таким образом без вмешательства в сенсор основной камеры видеотракта, следовательно, без ухудшения качества изображения на выходе видеотракта, за один шаг подстройки фокусировки удастся добиться сфокусированного состояния оптической системы на нужной точке сцены или на отслеживаемом объекте. В дальнейшем возможна интеграция методов машинного обучения для повышения устойчивости алгоритма сопоставления изображений, а также улучшения отслеживания движения объектов сцены.

Перед использованием системы необходимо выполнить калибровку стереопары технических камер. Она включает в себя определение параметров камер, таких как фокусное расстояние, число диафрагмы объектива технической камеры, а также определение взаимного расположения камер, то есть определение базы.

Также качество построения карты глубины существенно зависит от наличия текстурных особенностей на объектах сцены и уровня освещённости. Однородные поверхности, зеркальные и прозрачные объекты могут приводить к ошибкам сопоставления стереоизображений. В связи с этим актуальной является задача разработки адаптивных методов фильтрации и оценки достоверности значений глубин, позволяющих снизить вероятность ошибки.

Описанный метод может быть применён не только в системах автоматической фокусировки профессиональных кино- и телевизионных камер. Данный метод оценки глубины с помощью стереопары камер находит применение в роботизированных системах, устройствах виртуальной и дополненной реальности.

Заключение

В рамках данной работы предложен подход к решению задачи автоматической фокусировки мануальных объективов на основе стереоскопической оценки глубины сцены. В отличие от существующих методов автофокусировки, использующих активные дальнометры, встроенные фазовые сенсоры, или работающие на определении контраста датчики, разработанный метод основан на независимой системе стереозрения, не требует модификации сенсора основной камеры видеотракта. Это позволяет сохранить качество формируемого изображения и обеспечить предсказуемость работы системы фокусировки. Эти параметры являются критически важными для профессиональной киносъёмки или телевидения.

Научная новизна работы заключается в формировании методологии использования карты глубины, получаемой по стереоизображению для управления фокусировкой мануального объектива в реальном времени.

В работе предложена обобщённая структурная схема системы автофокусировки, описан математический аппарат вычисления расстояния до объектов сцены и определены формальные границы применимости метода с учётом параметров используемых в стереопаре камер.

Перспективы дальнейших исследований связаны с использованием алгоритмов компьютерного зрения и машинного обучения для повышения устойчивости системы к шумам, выбросам данных и неоднородностям сцены, а также к улучшению качества отслеживания движения объектов в пространстве.

Литература

1. Zhang Y., Liu L., Gong W., Yu H., Wang W., Zhao C., Wang P., Ueda T. Autofocus system and evaluation methodologies: a literature review // *Sensors and Materials*. 2018. Vol. 30, № 5, pp. 1165-1174. No. 5 (2018), pp. 1165-1174.
2. Herrmann C., Bowen R.S., Wadhwa N., Garg R., He Q., Barron J.T., Zabih R. Learning to autofocus // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020, pp. 2230-2239.
3. Kurita T., Kondo Y., Sun L., Sasaki T., Nitta S., Hashimoto Y., Muramatsu Y., Moriuchi Y. Revisiting disparity from dual-pixel images: physics-informed lightweight depth estimation // *arXiv:2411.04714v1 [cs.CV]*. 2024.
4. Juneau N. Low-cost LiDAR lens autofocus system for cinema cameras // *Inquiry Journal*. 2025. Spring issue.
5. Коняшов В.В., Сергеев А.С. (науч. рук. Федоров А.В.) Алгоритм построения карты глубины и оценка расстояния до объекта с помощью системы стереозрения // *Сборник тезисов докладов конгресса молодых ученых. Электронное издание. СПб: Университет ИТМО, [2023]. URL: <https://kmu.itmo.ru/digests/article/9887>.*
6. Huang S.K., Yang C.C., Shih K.T., Chen H.H. Using disparity information for stereo autofocus in 3-D photography // *Proceedings of IS&T International Symposium on Electronic Imaging: Digital Photography and Mobile Imaging XII*. 2016. P. DPMI-254.1–DPMI-254.6. DOI: 10.2352/ISSN.2470-1173.2016.18.DPMI-254.
7. Kamencay P., Breznan M., Jarina R., Lukac P., Zachariasova M. Improved depth map estimation from stereo images based on hybrid method // *Radioengineering*. 2012. Vol. 21, № 1, pp. 70-79.
8. Силантьева А.С., Власюк И.В. Анализ определения стереобазиса распространёнными методами для решения прикладных задач фотограмметрии // *Телекоммуникационные и вычислительные системы. Юбилейный сборник трудов тридцатого международного научно-технического форума*. 2022. С. 349-352.
9. Potashnikov A.M., Stroganova E.P., Vlasuyk I.V. Color contrast method based on subjective warmth and coldness // *T-Comm: Телекоммуникации и транспорт*. 2025. Vol. 19, No. 3, pp. 69-76. DOI 10.36724/2072-8735-2025-19-3-69-76. EDN BMCPUW.
10. Ivanchev V.V., Vlasuyk I.V., Stroganova E.P. Objective assessment of colours' warmth // *T-Comm: Телекоммуникации и транспорт*. 2024. Vol. 18, No. 1, pp. 44-50. DOI 10.36724/2072-8735-2024-18-1-44-50. EDN EKABXU.
11. Mozhaeva A., Vashenko E., Selivanov V. et al. Analysis of current video databases for quality assessment // *T-Comm: Телекоммуникации и транспорт*. 2022. Vol. 16, No. 2, pp. 48-56. DOI 10.36724/2072-8735-2022-16-2-48-56. EDN KRCZDN.
12. Романов С.Г., Власюк И.В. Методика расчета параметров анти-алайсинговых фильтров для коррекции спектральных характеристик в зависимости от используемых структур дискретизации массивов светофильтров // *T-Comm: Телекоммуникации и транспорт*. 2023. Т. 17, № 5. С. 4-13. DOI 10.36724/2072-8735-2023-17-5-4-13. EDN AKGLYF.