

Моделирование скрытых закономерностей в остаточных компонентах временных рядов для задач прогнозирования

А.В. Гуляев, С.В. Пивнева

Аннотация— В данной статье исследуется проблема повышения точности прогнозирования временных рядов за счет выявления и моделирования скрытых структурных закономерностей в их остаточных компонентах. Временные ряды, представляющие собой последовательности наблюдений, играют важнейшую роль в анализе данных в различных прикладных областях, однако их прогнозирование часто осложняется наличием сложных нелинейных зависимостей. Традиционно остаточная компонента, получаемая после декомпозиции ряда, интерпретируется как случайный шум, лишенный прогностической ценности. В первой части статьи проводится детальный обзор современных подходов, подчеркивающих растущий тренд на переосмысление данной парадигмы и акцент на извлечении полезных сигналов из остатков. Во второй части статьи особое внимание уделяется двум передовым гибридным архитектурам глубокого обучения: CNN-GRU и Transformer-GRU. Эти методы являются современными инструментами, которые позволяют комбинировать анализ локальных пространственных паттернов и глобальных временных зависимостей для моделирования сложных структур в остатках. Экспериментальное исследование, проведенное на реальных данных розничных продаж, социальной статистики и финансового ряда, подтвердило наличие структурных закономерностей в остаточных компонентах и установило, что эффективность гибридных архитектур детерминирована характером данных. Эта статья ставит своей целью содействие развитию методов анализа временных рядов, предоставляя инструмент для извлечения дополнительной информации, что напрямую влияет на качество построения прогнозов.

Ключевые слова— временные ряды, гибридные модели, остаточный компонент, CNN-GRU, Transformer-GRU, прогнозирование.

I. ВВЕДЕНИЕ

Современные исследования в области прогнозирования временных рядов демонстрируют растущий интерес к гибридным моделям, которые комбинируют методы декомпозиции данных с архитектурами глубокого обучения. Ключевой идеей,

Статья получена 30 мая 2026.

А.В. Гуляев — аспирант кафедры «Информационных технологий, искусственного интеллекта и общественно-социальных технологий цифрового общества» Российский государственный социальный университет (e-mail: Gulyaev81@gmail.com)

С.В. Пивнева — декан факультета политических и социальных технологий, кандидат педагогических наук, доцент. Российский государственный социальный университет (e-mail: tltswetlana@yandex.ru)

объединяющей эти работы, является признание того, что остаточные компоненты, традиционно интерпретируемые как шум, могут содержать скрытые структурные закономерности, использование которых позволяет существенно повысить точность прогноза [1].

Так, в работе [2] предложена гибридная модель ATES для прогнозирования продаж, где остатки после выделения линейной компоненты моделируются с помощью ансамблевой архитектуры GRU-CNN-LSTM. В исследовании [3] для прогнозирования цен на акции применяется частотная декомпозиция ряда методом CEEMD с последующей обработкой компонент архитектурой LSTM-GRU-Transformer. Близкая методологическая парадигма реализована в [4], где для прогнозирования цен на электроэнергию используется декомпозиция методом SSA с последующим моделированием компонент схемой CNN-GRU.

Целью настоящего исследования является проведение сравнительного анализа эффективности гибридных архитектур CNN-GRU и Transformer-GRU для прогнозирования остаточных компонентов временных рядов, выделенных после SSA-декомпозиции. Для достижения поставленной цели предполагается разработать методологическую схему, включающую этап декомпозиции исходных рядов методом SSA, адаптировать гибридные архитектуры CNN-GRU и Transformer-GRU для задачи прогнозирования остаточных компонентов, провести экспериментальное исследование на реальных данных различной природы и выполнить сравнительный анализ моделей на основе метрик MSE, MAPE и R^2 для выявления эмпирических закономерностей между типом данных и эффективностью гибридной архитектуры.

II. МЕТОДОЛОГИЯ

Для решения поставленной задачи была разработана методологическая схема, включающая три ключевых этапа:

1. Предварительная обработка данных.

Применен метод сингулярного спектрального анализа (SSA) [8] для декомпозиции временных рядов.

$$Z(t) = Z_{trend}(t) + Z_{cycle}(t) + Z_{res}(t)$$

где $Z_{cycle}(t)$ - периодическая компонента, $Z_{trend}(t)$ - компонента тренда, $Z_{res}(t)$ - остаточная компонента.

Выбор SSA обусловлен его эффективностью при работе с нестационарными рядами сложной структуры. Особое внимание уделялось контролю качества декомпозиции через тест Дики-Фуллера для остатков.

2. Архитектуры моделей.

Первая гибридная модель на основе архитектуры сверточной нейронной сети [9] сочетает CNN слои для выделения локальных пространственных паттернов и GRU-блоки для анализа временных зависимостей.

Временной ряд $Z_{res}(t)$ разбивается на два представления. Для CNN $X_{CNN} \in \mathbb{R}^{B \times T}$ где B – размер батча, а T – размер окна. Для GRU $X_{GRU} \in \mathbb{R}^{B \times T \times T}$

CNN-ветвь архитектуры предназначена для выделения локальных паттернов в остаточных компонентах. Входная последовательность X_{CNN} сначала преобразуется в тензорную форму, подходящую для обработки сверточными слоями.

Первый сверточный слой с функцией активации ReLU выделяет низкоуровневые характеристики:

$$Z^{(1)} = \text{ReLU}(W^{(1)} * X + b^{(1)})$$

За ним следует операция max-pooling для уменьшения размерности и повышения инвариантности к сдвигам:

$$P^{(1)} = \text{MaxPool}(Z^{(1)})$$

Второй сверточный слой обрабатывает полученные признаки для выявления более сложных зависимостей:

$$Z^{(2)} = \text{ReLU}(W^{(2)} * P^{(1)} + b^{(2)})$$

Полученные признаки преобразуются в вектор фиксированной размерности через операцию flatten, после чего передаются в полносвязные слои для формирования итогового представления H_{CNN} .

GRU-ветвь архитектуры служит для анализа временных зависимостей в остаточных компонентах. Обработка данных осуществляется двумя последовательными слоями GRU, которые преобразуют входную последовательность в последовательность скрытых состояний, сохраняющих информацию о временных зависимостях.

$$H_{GRU}^{(1)} = \text{GRU}(X_{GRU}; W_{gru1})$$

$$H_{GRU}^{(2)} = \text{GRU}(X_{GRU}; W_{gru2})$$

Полученные скрытые состояния затем проходят через три полносвязных слоя с функциями активации ReLU, которые последовательно преобразуют временные характеристики в фиксированное векторное представление H_{GRU} .

Объединение ветвей и построение прогноза осуществляется через конкатенацию векторных представлений, полученных от CNN- и GRU-ветвей.

$$X_{concat} = [X_{CNN}^{(7)} \parallel H_{GRU}^{(5)}]$$

$$\tilde{Z}_{res}(t) = \sigma_{selu}(W_{out} X_{concat} + b_{out})$$

Объединенный вектор проходит через финальный полносвязный слой с линейной функцией активации, который формирует итоговый прогноз остаточного компонента.

Предложенная архитектура демонстрирует эффективность при работе с зашумленными данными, поскольку CNN-компонент успешно фильтрует локальные шумы, а GRU-компонент сохраняет значимые временные зависимости.

Архитектура Transformer-GRU интегрирует механизм внимания для учета глобальных зависимостей и GRU-слои для временного моделирования. Данный подход демонстрирует преимущества при работе с длинными

последовательностями и рядами с выраженными долгосрочными зависимостями [10].

GRU-ветвь архитектуры реализована в виде двух последовательных GRU-слоев, которые преобразуют входную последовательность в скрытые состояния, сохраняющие информацию о временных зависимостях.

$$H_1 = \text{GRU}_1(X), \quad H_1 \in \mathbb{R}^{B \times T \times 64}$$

$$H_2 = \text{GRU}_2(H_1), \quad H_2 \in \mathbb{R}^{B \times 32}$$

Полученные состояния обрабатываются тремя полносвязными слоями для формирования векторного представления Y_{GRU} .

$$D_1 = \text{ReLU}(W_1 H_2 + b_1)$$

$$D_2 = \text{Flatten}(D_1)$$

$$Y_{GRU} = W_2 D_2 + b_2$$

Transformer-ветвь выполняет поэтапное преобразование входных данных. На первом этапе скалярные значения временного ряда проецируются в векторное пространство размерностью $d_{model}=64$ через обучаемый слой эмбеддингов.

$$X_{proj} = X \cdot W_p + b_p$$

К полученным эмбеддингам добавляется позиционное кодирование, вычисляемое по стандартной формуле синусоидальных функций, что позволяет модели учитывать порядок элементов в последовательности.

Далее данные проходят через четыре последовательных слоя трансформера. Каждый трансформерный слой содержит многоголовый механизм самовнимания с четырьмя головами, позволяющий модели одновременно анализировать различные типы зависимостей в данных. Механизм внимания вычисляет весовые коэффициенты для всех элементов последовательности, определяя их взаимное влияние.

После механизма внимания данные поступают в позиционно-зависимую полносвязную сеть, которая применяет нелинейное преобразование к каждому элементу последовательности независимо.

После обработки трансформерными слоями применяется сверточный слой для формирования итогового представления Y_{trans} :

$$C = \text{Conv1D}_{k=3}(X_{trans})$$

$$F = \text{Flatten}(C)$$

$$Y_{trans} = W_t F + b_t$$

Объединение ветвей осуществляется через конкатенацию векторных представлений Y_{GRU} и Y_{trans} с последующей проекцией через полносвязный слой для получения финального прогноза остаточного компонента.

$$Y_{concat} = [Y_{GRU} \parallel Y_{trans}], \quad Y_{concat} \in \mathbb{R}^{B \times 512}$$

Объединенный вектор последовательно проходит через серию полносвязных слоев которые уменьшают размерность данных. Финальный полносвязный слой с линейной активацией проецирует полученное представление в прогнозное значение остаточного компонента.

$$Z_{res} = D_{out} \circ D_4 \circ D_3 \circ D_2 \circ D_1([Y_{GRU}; Y_{trans}])$$

Такая схема позволяет эффективно комбинировать локальные временные зависимости, выявленные GRU, с

глобальными контекстными зависимостями, захваченными механизмом внимания Transformer.

отобраны три типа временных рядов, существенно различающихся по своей природе и характеристикам.

III. РЕЗУЛЬТАТЫ

Для проведения сравнительного анализа гибридных архитектур CNN-GRU и Transformer-GRU были

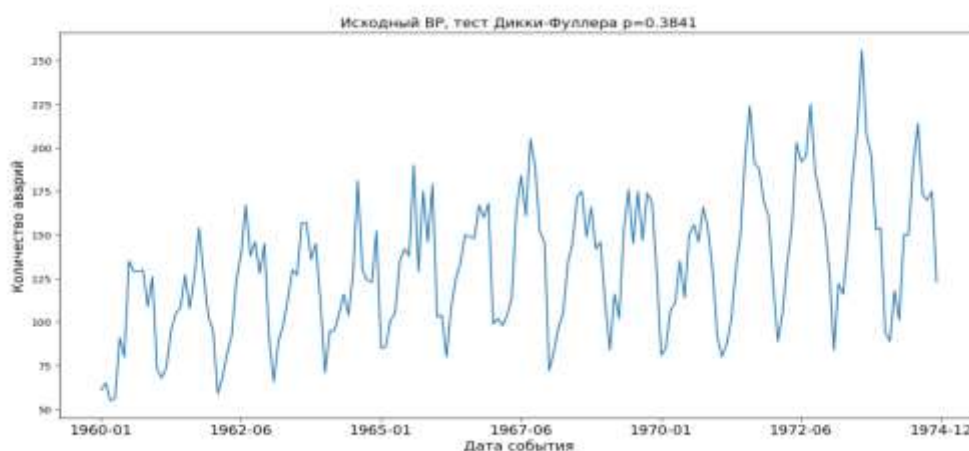


Рис. 1: Временной ряд дорожных происшествий в Онтарио (Канада) TRAFFAT

Первый исследуемый ряд представляет собой данные о ежедневных продажах продовольственных товаров. В выборку вошли 1000 различных SKU (товарных позиций) за период 1800 дней. Данный ряд характеризуется выраженной нестационарностью, наличием множественных сезонных компонент (недельная, месячная, годовая сезонность), а также значительным влиянием внешних факторов, таких как маркетинговые акции и ценовые колебания. Особенностью ряда является высокая вариативность остаточного компонента.

Второй анализируемый ряд - ежемесячная статистика дорожно-транспортных происшествий со смертельным исходом в провинции Онтарио (Канада) за период с января 1960 по декабрь 1974 года (TRAFFAT). Ряд содержит 180 наблюдений и отличается выраженной годовой сезонностью.

Третий ряд представляет собой ежедневные значения курса валютной пары USD/RUB за период с 1 февраля 2020 года по 5 февраля 2023 года. Особенностью данного ряда является экстремальная волатильность,

вызванная геополитическими событиями и макроэкономическими шоками. Амплитуда колебаний в отдельные периоды превышала 15% в сутки.

Для выделения остаточных компонентов во всех случаях применялся метод сингулярного спектрального анализа с последующим группированием компонент [12].

Для ряда продаж продовольственных товаров одного из SKU использовалось окно длиной $L=1/2$ длины ряда. 3 компоненты были отнесены к тренду, компоненты 6 к сезонности. Остальные компоненты были объединены в остаточный ряд.

При анализе данных о ДТП применялось окно $L=90$ месяцев. В результате декомпозиции было выделено 90 значимых компонент, из которых 3 отнесены к тренду, 8 к сезонности, а остальные компоненты составили остаточный ряд [13].

Для валютного ряда использовалось адаптивное окно $L=60$ дней. Трендовыми были признаны 4 компоненты, сезонными - 12, остальные компоненты образовали остаточный ряд.

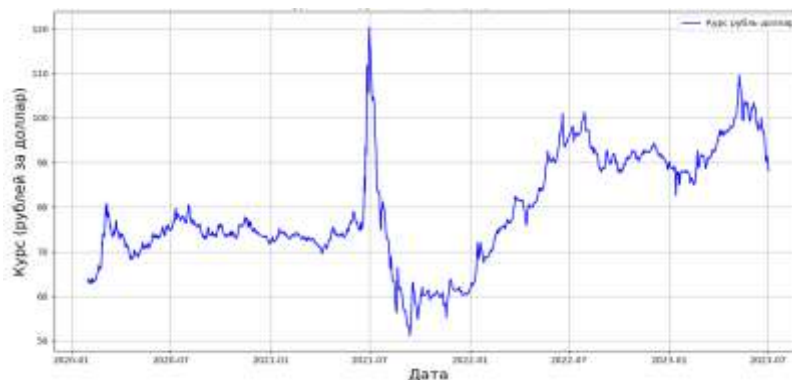


Рис. 2: Временной ряд дорожных происшествий в Онтарио (Канада) TRAFFAT

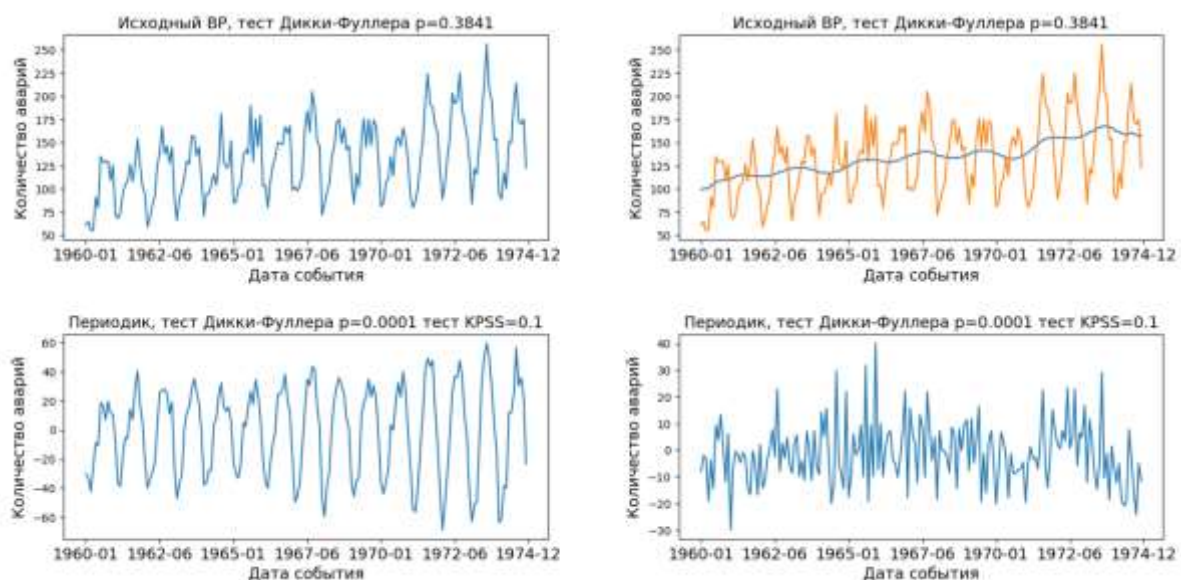


Рис. 3: Разложение временного ряда TRAFFAT с применением сингулярного спектрального анализа

Результаты экспериментального исследования, представленные в Таблице 1, демонстрируют выраженную зависимость эффективности моделей от характера данных.

Для временных рядов продаж продовольственных товаров модель CNN-GRU показывает стабильно более

точные прогнозы по сравнению с Transformer-GRU на всех протестированных горизонтах прогнозирования. Значения MSE и MAPE оказываются существенно ниже, а коэффициент детерминации R^2 - значительно выше для архитектуры CNN-GRU.

Таблица 1. Сравнительные результаты прогнозирования остаточных компонентов временных рядов

Характеристики ряда	Метрики	CNN-GRU	Transformer-GRU
Продажи продовольственных товаров Размер окна = 7, горизонт прогнозирования = 3	MSE MAPE R^2	1.54 57.8% 0.54	1.76 63.71% 0.48
Продажи продовольственных товаров Размер окна = 7, горизонт планирования = 5	MSE MAPE R^2	1.99 97.8% 0.34	2.57 122.25% 0.28
Дорожные происшествия (TRAFFAT) Размер окна = 24, горизонт планирования = 6	MSE MAPE R^2	175.78 109.1% 0.02	144.35 98.11% 0.15
Дорожные происшествия (TRAFFAT) Размер окна = 24, горизонт планирования = 3	MSE MAPE R^2	158.08 94.2% 0.04	124.82 88.15% 0.17
Курс USD/RUB Размер окна = 30, горизонт планирования = 5	MSE MAPE R^2	0.50 94.9% 0.57	0.85 141.90% 0.21
Курс USD/RUB Размер окна = 30, горизонт планирования = 3	MSE MAPE R^2	0.47 76.4% 0.62	0.78 131.37% 0.27

При работе с рядом дорожных происшествий (TRAFFAT) наблюдается противоположная картина. Архитектура Transformer-GRU демонстрирует более высокую точность прогнозирования, что проявляется в

меньших значениях MSE и MAPE, а также в более высоких показателях R^2 по сравнению с CNN-GRU.

Наиболее заметное преимущество CNN-GRU наблюдается на данных валютного курса USD/RUB. По всем метрикам качества архитектура CNN-GRU

показывает значительно лучшие результаты, особенно существенно превосходя Transformer-GRU по коэффициенту детерминации R^2 .

Анализ устойчивости моделей к увеличению горизонта прогнозирования показывает, что с ростом горизонта предсказания происходит ожидаемое снижение точности обеих моделей, однако соотношение их эффективности сохраняется для каждого типа данных.

На Рис. 4 и 5 представлены визуализации прогнозов остаточных компонентов для ряда продаж и валютного курса, построенные с использованием CNN-GRU архитектуры. Анализ графиков подтверждает способность модели точно отслеживать динамику остаточных компонентов различных типов временных рядов.

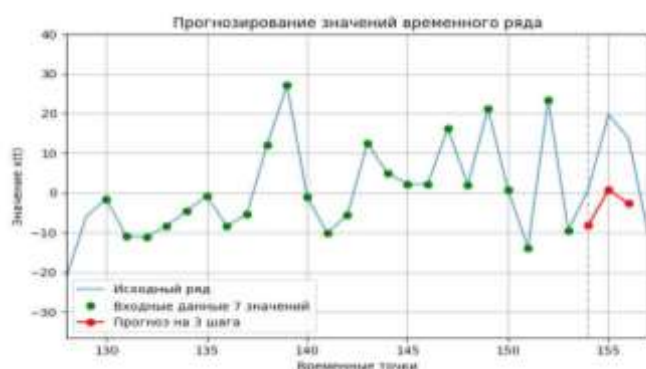


Рис. 4: Построение прогноза остаточного компонента с применением CNN-GRU архитектуры временного ряда продажи продовольственных товаров по выбранному SKU

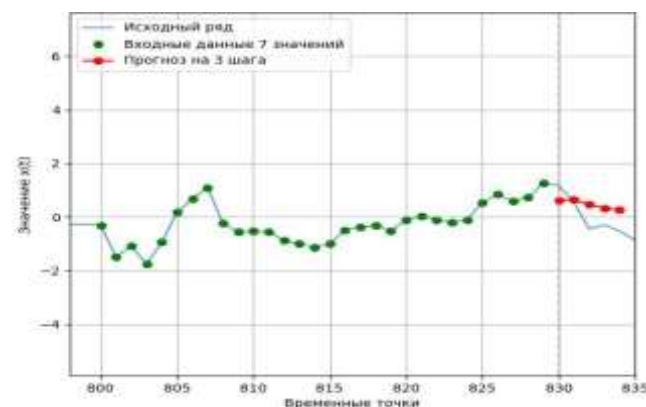


Рис. 5: Построение прогноза остаточного компонента временного ряда валютной пары с применением CNN-GRU архитектуры

Полученные результаты позволяют заключить, что выбор оптимальной архитектуры зависит от природы анализируемых данных. CNN-GRU демонстрирует лучшую производительность при работе с высокочастотными рядами (финансовые и товарные данные), в то время как Transformer-GRU более эффективен для рядов с выраженными долгосрочными зависимостями (социальные данные). Наблюдаемые закономерности подтверждаются всеми рассматриваемыми метриками (MSE, MAPE и R^2) и устойчивы к изменению горизонта прогнозирования.

IV. ВЫВОДЫ

Проведенное исследование было направлено на проверку гипотезы о возможности повышения точности прогнозирования временных рядов за счет выявления и моделирования скрытых структурных закономерностей в их остаточных компонентах. В рамках данной работы удалось достичь следующих результатов.

Экспериментально подтверждено наличие структурных закономерностей в остаточных компонентах после декомпозиции временных рядов методом SSA. Это открывает новые возможности для повышения точности прогнозов в прикладных задачах, где идеальная декомпозиция часто недостижима.

Установлена зависимость эффективности гибридных архитектур от характера исходных данных. Архитектура Transformer-GRU показала лучшие результаты на рядах с выраженными долгосрочными зависимостями, тогда как CNN-GRU проявила большую эффективность при работе с высокочастотными и зашумленными данными. Это позволяет рекомендовать выбор архитектуры в зависимости от специфики решаемой задачи.

Разработанный методический подход демонстрирует потенциал комбинирования методов декомпозиции с гибридными архитектурами глубокого обучения. Однако для подтверждения универсальности предложенного подхода требуются дальнейшие исследования на более широком круге задач.

Перспективные направления будущих исследований связаны с развитием адаптивных методов комбинирования архитектур CNN-GRU и Transformer-GRU в единой модели, способной автоматически выбирать оптимальную стратегию обработки в зависимости от характеристик временного ряда.

Полученные результаты свидетельствуют о перспективности дальнейшего развития методов анализа остаточных компонентов временных рядов и могут способствовать модернизации существующих подходов к прогнозированию в различных прикладных областях.

БИБЛИОГРАФИЯ

- [1] Гуляев А.В. / Применение вейвлет-преобразования и сингулярного спектрального анализа при декомпозиции временного ряда / Системы высокой доступности» №2 за 2024 г. с. 76-84 /DOI 10.18127/j20729472-202402-06
- [2] Deep-learning model using hybrid adaptive trend estimated series for modelling and forecasting sales / Md. I. A. Efat, P. Hajek, M. Z. Abedin [et al.] // Annals of Operations Research. – 2022. – DOI 10.1007/s10479-022-04838-6. – EDN HQYUOV.
- [3] Li, Ch. Stock Price Prediction Using a Frequency Decomposition Based GRU Transformer Neural Network / Ch. Li, G. Qian // Applied Sciences (Switzerland). – 2023. – Vol. 13, No. 1. – P. 222. – DOI 10.3390/app13010222. – EDN ZPWLHD.
- [4] L. Jiang, X. Zhang, P. Zhang, W. Pang and C. Qu, "Short-Term Tariff Prediction for High Penetration New Energy Electricity Market Based on Singular Spectrum Analysis and CNN-GRU Combination Model," 2024 International Conference on New Trends in Computational Intelligence (NTCI), Qingdao, China, 2024, pp. 166-171, doi: 10.1109/NTCI64025.2024.10776344.
- [5] A Arulmurugan, Et. Al. A Model for Predicting E-Commerce Product Returns Using Hybrid CNN-GRU / Et. Al. A Arulmurugan // International Journal on Recent and Innovation Trends in Computing and Communication. – 2023. – Vol. 11, No. 9. – P. 3615-3619. – DOI 10.17762/ijritcc.v11i9.9583. – EDN SJUPQN
- [6] Chen, Ya. Construction of a carbon neutral enterprise environmental performance assessment model based on transformer-GRU / Ya. Chen

- // *Frontiers in Ecology and Evolution*. – 2023. – Vol. 11. – DOI 10.3389/fevo.2023.1247644. – EDN ZTUIVX.
- [7] GRU- and Transformer-Based Periodicity Fusion Network for Traffic Forecasting / Ya. Zhang, Sh. Liu, P. Zhang, Bo. Li // *Electronics*. – 2023. – Vol. 12, No. 24. – P. 4988. – DOI 10.3390/electronics12244988. – EDN FZSTSE.
- [8] Поршнев, С. В. Исследование особенностей применения метода сингулярного спектрального анализа в задаче анализа и прогнозирования временных рядов : Монография / С. В. Поршнев, Ф. Рабайя. – Ульяновск : ИП Кеньшенская Виктория Валерьевна (издательство "Зебра"), 2016. – 167 с. – ISBN 978-5-9907847-8-9. – EDN VLYAZX.
- [9] Bogaba: a novel hybrid deep learning model for time series forecasting using sentiment analysis / U. Mandal, A. Chakraborty, Ph. Mahato [et al.] // *Indian Journal of Computer Science and Engineering*. – 2023. – Vol. 14, No. 1. – P. 130-145. – DOI 10.21817/indjcse/2023/v14i1/231401121. – EDN XEPYKM.
- [10] GRU–Transformer: A Novel Hybrid Model for Predicting Soil Moisture Content in Root Zones / W. Zheng, K. Zheng, L. Gao [et al.] // *Agronomy*. – 2024. – Vol. 14, No. 3. – P. 432. – DOI 10.3390/agronomy14030432. – EDN PUVAQJ.
- [11] Transformer-Bls: An Efficient Learning Algorithm Based on Multi-Head Attention Mechanism and Incremental Learning Algorithms / R. Fu, Ya. Wang, Ch. Jia, G. Sun // *Social Science Research Network*. – 2022. – DOI 10.2139/ssrn.4111793. – EDN UJNUKZ.
- [12] Swart S.B., Den Otter A.R., Lamothe C.J.C./Singular spectrum analysis as a data-driven approach to the analysis of motor adaptation time series/*Biomedical signal processing and control*/vol:71, 2022, p.103
- [13] Golyandina, N. Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing / N. Golyandina // *Wiley Interdisciplinary Reviews: Computational Statistics*. – 2020. – Vol. 12, No. 4. – P. e1487. – DOI 10.1002/wics.1487. – EDN GHQAN.

Modeling Latent Patterns in Time Series Residuals for Forecasting Purposes

A.V. Gulyaev, S.V. Pivneva

Abstract - This article investigates the problem of improving the accuracy of time series forecasting by identifying and modeling hidden structural patterns in their residual components. Time series, representing sequences of observations, play a crucial role in data analysis across various applied domains. However, their forecasting is often complicated by the presence of complex nonlinear dependencies. Traditionally, the residual component obtained after series decomposition is interpreted as random noise lacking predictive value. The first part of the article provides a detailed review of modern approaches, highlighting the growing trend of rethinking this paradigm and focusing on extracting useful signals from residuals. The second part of the article pays special attention to two advanced hybrid deep learning architectures: CNN-GRU and Transformer-GRU. These methods represent modern tools that combine the analysis of local spatial patterns and global temporal dependencies to model complex structures in residuals. Experimental research conducted on real-world data from retail sales, social statistics, and financial series confirmed the presence of structural patterns in residual components and established that the effectiveness of hybrid architectures is determined by the nature of the data. This article aims to contribute to the development of time series analysis methods by providing a tool for extracting additional information, which directly impacts the quality of forecast construction.

Keywords - time series, hybrid models, residual component, CNN-GRU, Transformer-GRU, forecasting.

REFERENCES

- [1] Gulyaev A.V. Application of wavelet transform and singular spectrum analysis for time series decomposition. *Systems of High Availability*. 2024, no. 2, pp. 76–84. DOI: 10.18127/j20729472-202402-06.
- [2] Deep-learning model using hybrid adaptive trend estimated series for modelling and forecasting sales / Md. I. A. Eftekhar, P. Hajek, M. Z. Abedin [et al.] // Annals of Operations Research. – 2022. – DOI 10.1007/s10479-022-04838-6. – EDN HQYUOV.
- [3] Li, Ch. Stock Price Prediction Using a Frequency Decomposition Based GRU Transformer Neural Network / Ch. Li, G. Qian // Applied Sciences (Switzerland). – 2023. – Vol. 13, No. 1. – P. 222. – DOI 10.3390/app13010222. – EDN ZPWLHD.
- [4] L. Jiang, X. Zhang, P. Zhang, W. Pang and C. Qu, "Short-Term Tariff Prediction for High Penetration New Energy Electricity Market Based on Singular Spectrum Analysis and CNN-GRU Combination Model," 2024 International Conference on New Trends in Computational Intelligence (NTCI), Qingdao, China, 2024, pp. 166-171, doi: 10.1109/NTCI64025.2024.10776344.
- [5] Arulmurugan A. et al. A Model for Predicting E-Commerce Product Returns Using Hybrid CNN-GRU. *International Journal on Recent and Innovation Trends in Computing and Communication*. 2023, vol. 11, no. 9, pp. 3615–3619. DOI: 10.17762/ijritcc.v11i9.9583. EDN: SJUPQN.
- [6] Chen Ya. Construction of a carbon neutral enterprise environmental performance assessment model based on transformer-GRU. *Frontiers in Ecology and Evolution*. 2023, vol. 11. DOI: 10.3389/fevo.2023.1247644. EDN: ZTUIVX.
- [7] Zhang Ya., Liu Sh., Zhang P., Li Bo. GRU- and Transformer-Based Periodicity Fusion Network for Traffic Forecasting. *Electronics*. 2023, vol. 12, no. 24, p. 4988. DOI: 10.3390/electronics12244988. EDN: FZSTSE.
- [8] Porshnev S.V., Rabaya F. *Study of Singular Spectrum Analysis Application Features in Time Series Analysis and Forecasting: A Monograph*. Ulyanovsk: Zebra, 2016. 167 p. ISBN 978-5-9907847-8-9. EDN: VLYAZX.
- [9] Mandal U., Chakraborty A., Mahato Ph. et al. Bogaba: a novel hybrid deep learning model for time series forecasting using sentiment analysis. *Indian Journal of Computer Science and Engineering*. 2023, vol. 14, no. 1, pp. 130–145. DOI: 10.21817/indjcse/2023/v14i1/231401121. EDN: XEPYKM.
- [10] Zheng W., Zheng K., Gao L. et al. GRU-Transformer: A Novel Hybrid Model for Predicting Soil Moisture Content in Root Zones. *Agronomy*. 2024, vol. 14, no. 3, p. 432. DOI: 10.3390/agronomy14030432. EDN: PUVAQJ.
- [11] Fu R., Wang Ya., Jia Ch., Sun G. Transformer-Bls: An Efficient Learning Algorithm Based on Multi-Head Attention Mechanism and Incremental Learning Algorithms. *Social Science Research Network*. 2022. DOI: 10.2139/ssrn.4111793. EDN: UJNUKZ.
- [12] Swart S.B., Den Otter A.R., Lamothe C.J.C. Singular spectrum analysis as a data-driven approach to the analysis of motor adaptation time series. *Biomedical Signal Processing and Control*. 2022, vol. 71, p. 103.
- [13] Golyandina N. Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2020, vol. 12, no. 4, p. e1487. DOI: 10.1002/wics.1487. EDN: GHQAN.