

УДК 519.246.8
DOI 10.17513/snt.39974

ПРЕДИКТИВНЫЙ АНАЛИЗ ПОКАЗАТЕЛЕЙ КРИТИЧЕСКИХ СОБЫТИЙ С ПОМОЩЬЮ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ С ТРАНСФОРМЕРОМ

Полежаев М.Н., Финогеев А.А.

ФГБОУ ВО «Пензенский государственный университет», Пенза,
e-mail: max.polezhaev.01@mail.ru, fanton3@yandex.ru

Аннотация. Цель работы представляет собой предиктивный и сравнительный анализы временных рядов показателей аварийности и факторов влияния для обнаружения корреляций между их сегментами в прошлом и прогнозированием появления аналогичных паттернов и их сочетаний, которые приводят к рискам возникновения критических событий, в будущем. Анализ показателей аварийности и причин их возникновения может быть выполнен на базе исследования данных, представленных в виде набора временных рядов. В статье рассматриваются вопросы применения методов глубокого обучения и нейронных сетей для прогнозирования и сравнительного анализа будущих сочетаний паттернов временных рядов. В качестве метода предиктивного анализа предлагается использовать подход с использованием модели рекуррентной нейронной сети, дополненной архитектурой трансформера. Механизм внимания трансформера позволяет эффективно прогнозировать длинные временные ряды показателей аварийности и факторов влияния. Временные ряды формируются по данным, которые извлекаются с фоторадарных комплексов фото- и видеофиксации в дорожно-транспортной среде. Результатами исследований является модель рекуррентной нейронной сети с трансформером, методика ее синтеза и глубокого обучения на примере больших данных о зафиксированных проездах транспортных средств и нарушениях ими скоростного режима на участках автомагистралей.

Ключевые слова: безопасность дорожного движения, рекуррентная нейронная сеть, трансформер, механизм внимания, временной ряд, прогнозирование временного ряда, показатели аварийности

Исследование выполнено при поддержке гранта Российского научного фонда (проект № 20-71-10087).

PREDICTIVE ANALYSIS OF CRITICAL EVENTS INDICATORS USING A RECURRENT NEURAL NETWORK WITH A TRANSFORMER

Polezhaev M.N., Finogeev A.A.

Penza State University, Penza, e-mail: max.polezhaev.01@mail.ru, fanton3@yandex.ru

Annotation. The purpose of the work is a predictive and comparative analysis of time series of accident rates and influencing factors to detect correlations between their segments in the past and predict the emergence of similar patterns and combinations that lead to risks of critical events in the future. Analysis of accident rates and the causes of their occurrence can be performed on the basis of a study of data presented in the form of a set of time series. The article discusses the use of deep learning methods and neural networks for forecasting and comparative analysis of future combinations of time series patterns. As a method of predictive analysis, it is proposed to use an approach using a recurrent neural network model, complemented by a transformer architecture. The transformer attention mechanism can effectively predict long time series of accident rates and influencing factors. Time series are formed based on data that is extracted from photo radar photo and video recording systems in the road transport environment. The results of the research are a model of a recurrent neural network with a transformer, a methodology for its synthesis and deep learning using the example of big data on recorded passages of vehicles and violations of speed limits on sections of highways.

Keywords: road safety, recurrent neural network, transformer, attention mechanism, time series, time series forecasting, accident rates

The study was supported by a grant from the Russian Science Foundation (project no. 20-71-10087).

Во всем мире развитие дорожно-транспортной инфраструктуры вместе с ростом количества транспортных средств в последние годы ставит на первое место необходимость обеспечения безопасности на дорогах. Дорожно-транспортные происшествия (ДТП) являются серьезной проблемой, которая влечет за собой материальные и финансовые потери, но, что более важно, связаны с причинением вреда здоровью человека и потерями человеческих жизней. Одним из аспектов повышения эффективности мероприятий по обеспечению безопасности

дорожного движения является выявление и анализ влияния различных факторов на показатели аварийности. Снижение числа ДТП посредством прогноза рисков инцидентов в зависимости от сочетания и изменения факторов влияния является целью проактивного управления безопасностью на дорогах. Изменения значений показателей аварийности и факторов влияния можно представить в виде временных рядов. Тогда задача прогноза рисков ДТП по сочетанию факторов и значениям их показателей сводится к выявлению схожих паттернов

во временных рядах методом сравнительного анализа и прогнозированию появления похожих паттернов и их сочетаний. Традиционные методы анализа и прогнозирования временных рядов, основанные на статистических и математических моделях, имеют ограничения в отношении работы с большими данными и скрытыми закономерностями, которые влияют на динамику показателей аварийности. В качестве эффективного инструментария для анализа временных рядов и их прогнозирования последнее время используются методы интеллектуального анализа и глубокого машинного обучения нейронных сетей.

Краткий обзор состояния исследований

Для поиска и определения «скрытых» закономерностей в больших данных применяются методы статистического, сравнительного и интеллектуального анализа, модели нейронных сетей и методы машинного обучения. Статистический анализ временных рядов применяется для обнаружения пропусков, аномальных отклонений, трендов, сезонных колебаний, случайных отклонений, для расчета различных показателей. Для прогнозирования рядов применяются методы авторегрессионного интегрированного скользящего среднего (ARIMA), модель экспоненциального сглаживания, структурные модели, модель дерева регрессии и т.д. В последнее время для анализа и прогноза временных рядов используются методы глубокого машинного обучения с использованием моделей нейронных сетей обобщенной регрессии (GRNN) [1], нейронных сетей эхо состояний (ESN) [2], рекуррентных нейронных сетей (RNN) [3], сверточных нейронных сетей (CNN) [4] и трансформеров (TSF) [5]. Сети RNN эффективны для извлечения скрытой информации из временных рядов. Недостатком классической модели RNN является проблема затухающего градиента при анализе больших данных. Для решения проблемы разработаны модели с долгой краткосрочной памятью (LSTM) [6] и с управляемым рекуррентным вентилем (GRU) [7]. Однако для исследования долгосрочных корреляций в спектре временных рядов необходима более сложная и устойчивая модель. В последнее время наблюдается тенденция применения моделей трансформеров [8, 9] в различных задачах искусственного интеллекта. Проблема трансформеров заключается в ограниченных возможностях для представления последовательных отношений. Поэтому интересным подходом представляется интеграция модели RNN с моделью трансформера [10].

Целью исследования является развитие проактивного подхода к мониторингу и анализу возможных причин критических событий на объектах распределенной инфраструктуры на основе предиктивного анализа больших данных для выявления корреляционных связей между временными рядами характеристик событий и показателей возможных факторов влияния. Результатом проведенной работы является прогностическая оценка рисков критических событий на различных участках в зависимости от оценки состояния дорожно-транспортной инфраструктуры в совокупности с другими факторами. Задачи исследования связаны с разработкой нового подхода к прогнозированию временных рядов показателей инцидентов и оценки рисков влияния совокупности факторов и метода определения прогнозных моментов времени с превышением допустимого уровня рисков возникновения критических событий на участках дорог. Объектами исследований являются распределенная дорожно-транспортная инфраструктура, участники дорожного движения и эксплуатационное состояние дорожной инфраструктуры.

Материалы и методы исследования

Информация о транспортных и пешеходных потоках, ДТП и нарушениях правил дорожного движения (ПДД) собирается различными сенсорными устройствами, включая камеры видеонаблюдения, фоторадарные комплексы фото- и видеофиксации, детекторы движения и т.д. Данные о возможных факторах влияния, таких как погодные условия, видимость, освещенность, эксплуатационное состояние дорог, можно получить с дорожных погодных станций, детекторов состояния дорожного покрытия, камер видеонаблюдения, мобильных дорожных лабораторий, с метеорологических сервисов и т.д. Динамика изменения показателей дорожных инцидентов и возможных факторов влияния может быть представлена в виде набора временных рядов. Исследование сегментов временных рядов разных показателей в одинаковые временные интервалы позволяет обнаружить корреляции между фактами дорожных инцидентов и факторами влияния. Для уточнения корреляционных связей применяется метод сравнительного анализа паттернов рядов результирующих показателей аварийности и нарушений с паттернами рядов показателей факторов влияния. Целью анализа является выделение схожих и повторяющихся паттернов при попарном сравнении временных рядов. Примерами водных данных являются характеристики транспортных

потоков, данные о проездах и скоростях транспортных средств (ТС), данные о нарушениях и ДТП, характеристики метеорологических условий (температура, давление, видимость, освещенность, сила ветра, осадки и т.п.), показатели эксплуатационного состояния дорожного покрытия. В качестве результирующих показателей выбраны количество ДТП и нарушений ПДД на участках дорог. Для исследования корреляций между рядами применялись ранее разработанные методы статистического анализа [11], бенчмаркинг-анализа [12], интеллекту-

ального анализа и нечеткой логики [13], предиктивного анализа временных рядов [14].

Метод предиктивного анализа спектра временных рядов, синтезированных по данным показателей аварийности, нарушений и факторов влияния в дорожной среде, базируется на технологии глубокого машинного обучения для интегрированной модели рекуррентной нейронной сети (RNN) и модели трансформера в качестве механизма внимания. Эффективность такой модели подтверждается при работе с большими данными при увеличении времени обучения.

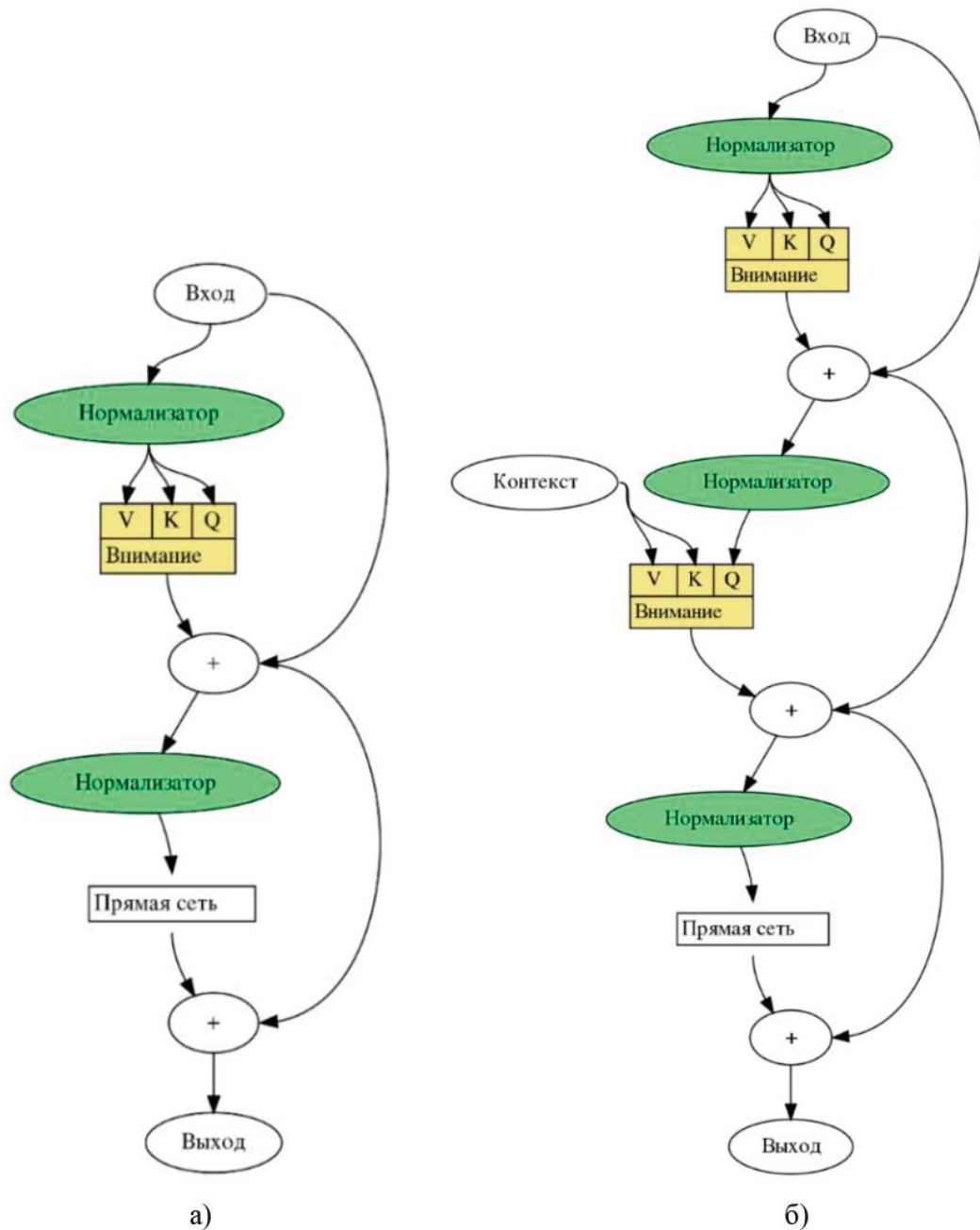


Рис. 1. Кодировочный а) и декодирующий б) слои трансформера

Для решения задачи прогнозирования необходимы подготовленные ряды данных. Процедура подготовки включает сбор, фильтрацию, нормализацию данных и представление их последовательности в виде временного ряда. Особое внимание уделяется заполнению пропусков в данных. Пропуски возникают в случае сбоя в работе устройства или сетевого оборудования. Также на этапе подготовки выполняется обнаружение аномалий и дубликатов. Работа с аномалиями и дубликатами сводится к восстановлению пропусков в данных, который основан на алгоритме k – ближайших соседей. Далее синтезируется модель RNN с архитектурой трансформера и выполняется настройка кодирующих и декодирующих слоев трансформера (рис. 1).

Следующим этапом является обучение модели на подготовленных данных. Данные делятся на обучающую и тестовую выборки, определяется функция потерь и выбирается оптимизатор. В ходе обучения настраиваются весовые коэффициенты на каждом слое методом градиентного спуска с целью минимизации функции потерь. После обучения выполняется оценка модели на тестовой выборке. В завершение выполняется визуализация прогнозируемого сегмента и сравнение с фактическим рядом. Если результаты прогноза не устраивают, то выполняется корректировка модели для повышения точности. Настройка включает изменение базовых параметров и структура модели (весовых коэффициентов, функции потерь, числа слоев, количества нейронов в слое).

Результаты исследования и их обсуждение

Для синтеза временных рядов получены данные с фоторадарных комплексов фото и видеофиксации дорожных событий, таких как проезды ТС, нарушения ПДД, показатели ДТП. Информация о показателях

относится к большим данным. Например, число проездов за месяц по участкам дорог Пензенской области, зафиксированных шестью устройствами, составило 40843 ТС с 6856 нарушениями ограничений скорости (табл. 1).

Более интенсивный трафик наблюдается на трассах федерального значения. Например, данные о проездах, скоростях и нарушениях скоростного режима представляют более миллиона записей за 1 месяц работы 5 комплексов на участках дороги М5 «Урал» в Самарской области (1036, 1043, 1044, 1107, 1154 км), включая 2713 фактов превышения более чем на 20 км/ч и 214 фактов превышения на 40 км/ч (рис. 2).

Предварительная обработка выполняется по каждому показателю с целью представления в виде табличных данных (табл. 2).

По данным синтезируется временной ряд для обучения нейронной сети (рис. 3).

Для синтеза временного ряда факторов влияния выгружены метеорологические данные с сайта <https://www.worldweatheronline.com>. Временные ряды синтезированы для показателей: средней температуры (avgtempC), температуры точки росы (DewPointC), температуры охлаждения ветром (WindChillC), высоты снежного покрова (totalSnow_cm), длительности светлого времени (sunHour), скорости ветра (windspeedKmph), видимости (visibility) и влажности (humidity).

Методика предиктивного анализа включает следующие этапы:

1. Подготовка данных. На данном шаге выполняется предварительная обработка табличных данных, полученных в результате выгрузки с комплексов фото- и видеофиксации с целью синтеза наборов для временных рядов прогнозируемых показателей, представляющих собой сумму синусоиды и случайных шумов.

Таблица 1

Пример сводных данных о проездах и нарушениях

Номер устройства	Дата (месяц, год)	Время работы комплекса	Фактическая фиксация нарушений комплексом %	Кол-во проездов ТС	Кол-во нарушений	% количества нарушений от числа проездов
KD0168	июн. 23	77,36	77,58	13362	3575	26,7549768
KD0171	июн. 23	98,33	98,54	8009	426	5,319016107
KD0172	июн. 23	97,99	98,19	7824	986	12,60224949
KD0173	июн. 23	4,17	100	113	21	18,5840708
KD0175	июн. 23	97,99	98,19	6578	783	11,90331408
KD0176	июн. 23	54,72	54,76	4957	1065	21,48476901

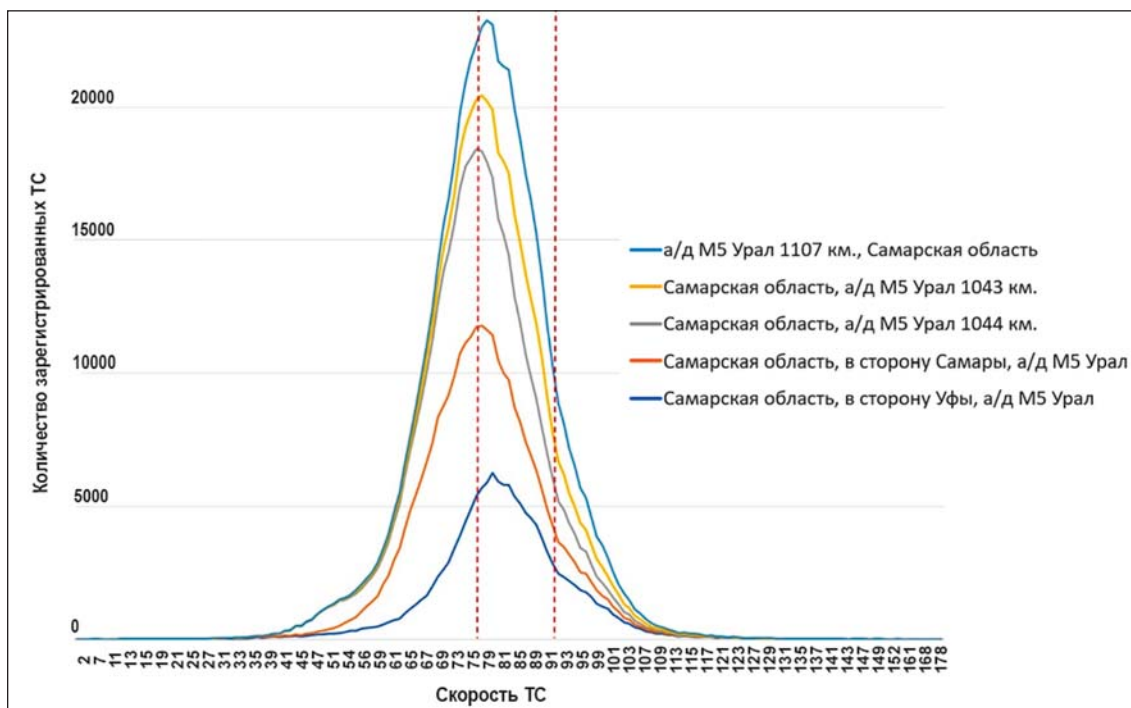


Рис. 2. Результаты кластеризации ТС по скоростям и нарушениям на пяти участках

Таблица 2

Фрагмент с данными о скоростях ТС

Дата	Время	Ограничение	Скорость	Номер ТС	Место фиксации
16.05.2023	11:27	60	50	о722рс58	Пензенская область, р.п. Колышлей. В сторону Пензы, 55-й км дороги г. Тамбов – г. Пенза – г. Сердобск – р.п. Беково
16.05.2023	11:27	60	60	м789вн163	Пензенская область, р.п. Колышлей. В сторону Пензы, 55-й км дороги г. Тамбов – г. Пенза – г. Сердобск – р.п. Беково
16.05.2023	11:27	60	65	р736хк58	Пензенская область, р.п. Колышлей. В сторону Беково, 55-й км дороги г. Тамбов – г. Пенза – г. Сердобск – р.п. Беково
16.05.2023	11:26	60	71	о229тм58	Пензенская область, р.п. Колышлей. В сторону Беково, 55-й км дороги г. Тамбов – г. Пенза – г. Сердобск – р.п. Беково



Рис. 3. Фрагмент временного ряда показателя скорости

2. Формирование выборок данных. Подготовленные наборы данных разбиваются на обучающую и тестовую выборки в соотношении 70% на 30%. Также методом скользящего окна создаются сегменты ряда для обучающих выборок.

3. Синтез рекуррентной нейронной сети Basic RNN с помощью пакета TensorFlow и библиотеки Keras. Для работы с нашими данными выбрана модель LSTM.

4. Интеграция с трансформером. К модели добавляются слои трансформера с использованием функции `layer_self_attention` для настройки способности модели к анализу долгосрочных зависимостей в больших данных.

5. Компиляция модели рекуррентной сети с трансформером. В процессе компиляции выполняется установка функции потерь (loss function) в виде среднеквадратичного отклонения (Mean Squared Error) и выбирается оптимизатор (Adam) для ее обучения.

6. Обучение модели. В процессе обучения задается число эпох обучения и размер пакета (batch size).

7. Прогнозирование временного ряда с использованием функции `predict`, визуализация результатов и сравнение с тестовым рядом для оценки точности.

Заключение

Растущее использование сенсорных устройств и средств мобильной связи привело к росту больших объемов сенсорных данных различного вида, которые необходимо обрабатывать для обнаружения и анализа «скрытых» зависимостей с целью прогнозирования рисков дорожных инцидентов и оценки влияния факторов их появления. В процессе исследований разработана модель рекуррентной нейронной сети с архитектурой трансформера. Благодаря слою внимания в трансформере выполняется параллельная обработка информации, что позволяет после обучения модели выявлять общие паттерны и прогнозировать временные ряды показателей инцидентов. Основным ограничением модели трансформера является необходимость большой обучающей выборки, что делает ее более затратной в плане вычислений. Также недостатком является относительно большое время обучения модели и число эпох, выделенных на обучение. Поэтому целью в ходе дальнейших исследований предполагается

доработать модель в плане компенсации данных ограничений.

Список литературы

1. Martínez Francisco, Maria P. Frias, M. Dolores Pérez-Godoy, Antonio Jesús Rivera Rivas. Time Series Forecasting by Generalized Regression Neural Networks Trained with Multiple Series // *IEEE Access* PP. 2022. Vol. 10. P. 3275–3283. DOI: 10.1109/ACCESS.2022.3140377.
2. Kim T., King B.R. Time series prediction using deep echo state networks // *Neural Comput & Applic* 32. 2020. Vol. 32. P. 17769–17787. DOI: 10.1007/s00521-020-04948-x.
3. Gábor Petneházi. Recurrent neural networks for time series forecasting // *Computer Science, Machine Learning*. 2019. DOI: 10.48550/arXiv.1901.00069.
4. Shaojie Bai, J Zico Kolter, Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling // *Computer Science, Machine Learning*. 2018. DOI: 10.48550/arXiv:1803.01271.
5. Li Z., Zhang X., Dong Z. TSF-transformer: a time series forecasting model for exhaust gas emission using transformer. *Appl Intell* 53. 2023. P. 17211–17225. DOI: 10.1007/s10489-022-04326-1.
6. Zhou X., Hu Y., Liang W., Ma J., and Jin Q. Variational LSTM enhanced anomaly detection for industrial big data // *IEEE Trans. Ind. Informat.* Vol. 17, Is. 5. 2021. P. 3469–3477.
7. Chung J., Gülçehre Ç., Cho K., Bengio Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling // *Computer Science, Neural and Evolutionary Computing*. 2014. DOI: 10.48550/arXiv.1412.3555.
8. Qingsong Wen, Tian Zhou, Chaoli Zhang, Weiqi Chen, Ziqing Ma, Junchi Yan, Liang Sun. Transformers in time series: A survey // *Computer Science, Machine Learning*. 2022. DOI: 10.48550/arXiv:2202.07125.
9. Lee Sangwon, Hong Junho, Liu Ling, Choi Wonik. TS-Fastformer: Fast Transformer for Time-Series Forecasting // *ACM Transactions on Intelligent Systems and Technology*. 2024. Vol. 15, Is. 2. № 24. P. 1–20.
10. Sakatani Y. Combining RNN with Transformer for Modeling Multi-Leg Trips // *ACM WSDM WebTour* 2021. 2021. P. 50–52.
11. Anton Finogeev, Michael Deev, Alexey Finogeev, Ilja Kolesnikoff. Proactive Big Data Analysis for Traffic Accident Prediction // 2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA), Sydney, Australia. 2020. P. 1–9.
12. Leyla Gamidullaeva, Alexey Finogeev, Sergey Vasin, Vardan Mkrttchian, Artiom Berezin. The use of benchmarking tool to improve efficiency of company's innovation activities in the conditions of digital economy // *International Journal of Process Management and Benchmarking*. 2021. Vol. 11, Is. 2. P. 151–177.
13. Alexey Finogeev, Michael Deev, Danila Parygin, Anton Finogeev. Intelligent SDN Architecture with Fuzzy Neural Network and Blockchain for Monitoring Critical Events // *Applied Artificial Intelligence*. 2022. Vol. 36, Is. 1. DOI: 10.1080/08839514.2022.2145634.
14. Колесников И.Н., Финогеев А.Г. Проактивный мониторинг событий на основе предиктивного анализа временных рядов // *Модели, системы, сети в экономике, технике, природе и обществе*. 2020. № 1 (33). С. 111–125. DOI: 10.21685/2227-8486-2020-1-9.