

# Прогнозирование трендовой, сезонной и нерегулярной составляющих среднего уровня аварийности на дорогах общего пользования с помощью рекуррентных нейронных сетей

А. Ю. Абалаяев, Л. В. Грунская

Владимирский государственный университет имени А.Г. Столетова и Н.Г. Столетова

ice\_aya@bk.ru, grunsk@vlsu.ru

**Аннотация.** В работе исследуется возможность построения прогноза среднего уровня аварийности на дорогах общего пользования с помощью рекуррентных нейронных сетей. Для увеличения точности прогноза используется разложение временного ряда, соответствующего среднему уровню аварийности, на трендовую, сезонную и нерегулярную компоненты. Приводится описание архитектуры каскадной нейронной сети, осуществляющей прогнозирование как на основе аддитивной, так и на основе мультипликативной декомпозиции временного ряда. Проводится сравнительная оценка точности полученного прогноза.

**Ключевые слова:** рекуррентные нейронные сети, аварийность, прогнозирование, временной ряд

## I. ВВЕДЕНИЕ

Несмотря на постоянный рост автомобильного парка, так, в 2024 году в РФ было реализовано порядка 1,59 млн. новых легковых автомобилей, развитие дорожной сети и инфраструктуры, а также оптимизация транспортных потоков, позволяет добиться ежегодного снижения числа дорожно-транспортных происшествий.

Однако, несмотря на это, автомобиль остается средством повышенной опасности, и, в случае попадания в аварийную ситуацию, несет угрозу здоровью, а иногда и жизни, участников дорожного движения.

Необходимость своевременного реагирования на происшествия на дорогах общего пользования и обеспечение экстренных служб достаточным количеством средств спасения и дежурных экипажей позволяет сделать вывод об актуальности задачи прогнозирования уровня аварийности на дорогах общего пользования. Решение этой задачи с помощью статистических методов не всегда позволяет обеспечить высокую точность. Использование технологий искусственного интеллекта и методов машинного обучения позволяет увеличить точность получаемого прогноза [1].

## II. МЕТОДЫ ИССЛЕДОВАНИЯ

Использование рекуррентных нейронных сетей для прогнозирования среднего уровня аварийности на дорогах общего пользования, особенно, с учетом дополнительных факторов, влияющих на психофизиологическое состояние человека, показало свою эффективность [2]. Тем не менее, как показано в [3], временной ряд суточного уровня аварийности на региональных дорогах общего пользования [4]

отличается определяющим влиянием случайной составляющей. Для определения степени влияния трендовой, сезонной и случайной составляющих среднего уровня аварийности, проводилась аддитивная декомпозиция временного ряда [5]:

$$y(t) = T(t) + S(t) + R(t),$$

где  $t$  – независимая переменная, время;  $T$  – трендовая составляющая;  $S$  – сезонная составляющая;  $R$  – случайная составляющая.

Трендовая составляющая вычисляется как скользящее среднее

$$T(t) = \frac{1}{m} \sum_{j=-k}^k y(t+j),$$

где  $m = 2k + 1$  – порядок скользящей средней.

Величина  $m$  выбирается исходя из длительности сезонной компоненты  $p$ . Если  $p$  – чётное, то  $m = 2p$ , а если нечётное – то  $m = p$ .

Сезонная компонента вычисляется как последовательность средних величин ряда  $y(t) - T(t)$  за период  $p$ . Сезонная составляющая корректируется таким образом, чтобы  $\sum S(t) = 0$ . Случайная составляющая вычисляется как

$$R(t) = y(t) - T(t) - S(t).$$

Для аддитивной декомпозиции использовалась выборка данных о суточной аварийности на дорогах общего пользования за период 2001 – 2010 гг. [4]. Общий объём выборки составил 3652 отчета.

Декомпозиции подвергались нормализованные данные, что позволяет упростить сравнение результатов обучения нейронных сетей по прогнозированию составляющих временного ряда и дальнейшую их обработку. Аддитивная декомпозиция проводилась для длительности сезонной компоненты  $p = 30$  суток. Такой период сезонности соответствует выбранной длине входной последовательности для обучения рекуррентной нейронной сети (РНС). В результате декомпозиции получены последовательности для трендовой, сезонной и случайной составляющих, состоящих из 3600 отсчетов каждая. Сезонная составляющая, с периодом 30 суток, для обеспечения единообразия в обработке данных, была распространена на полную продолжительность изучаемого периода.

Полученные временные ряды разбивались на последовательности из 30 входных и 7 выходных значений РНС. Для обучающей выборки использовались четыре пятых полученных последовательностей, для контрольной – оставшаяся одна пятая последовательностей.

Для решения задачи прогнозирования трендовой, сезонной и случайной составляющих использовалась архитектура РНС, показавшая наилучшие результаты при прогнозировании среднего уровня аварийности на дорогах общего пользования с учетом влияния солнечной активности на психофизиологическое состояние человека [3, 6]. РНС состоит из двух рекуррентных слоев LSTM-нейронов, слоя нормализации (*BatchNormalization*), двух скрытых слоев обратного распространения (*BackPropagation*) и выходного слоя. Для исключения переобучения между скрытыми слоями добавлен слой отбрасывания (*Dropout*). Параметр слоя отбрасывания – вероятность исключения синаптического веса из обучения нейронной сети, не должен превышать 0,1. В противном случае происходит существенная потеря точности обучения и валидации РНС.

Для прогнозирования среднего уровня аварийности используются каскадная нейронная сеть, состоящая из трёх РНС, описанных выше. Каждая из трех подсетей отвечает за прогнозирование соответствующей составляющей временного ряда *T*, *S* и *R*. Подсети работают параллельно, независимо друг от друга. На вход каскадной нейросети подается набор из трех векторов, размером 30 элементов. Каждый вектор соответствует трендовой, сезонной и случайной составляющей временного ряда. Результаты работы РНС – векторы прогнозных значений составляющих временного ряда на 7 суток вперед, суммируются почленно. Результатом работы каскадной сети, т. о., является прогноз среднего уровня аварийности на ближайшую неделю.

### III. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

Результаты аддитивной декомпозиции представлены на рис.1. Продолжительность трендовой и случайной составляющих, представленных на рисунке, равна 3600 суток, сезонной – 45 суток.

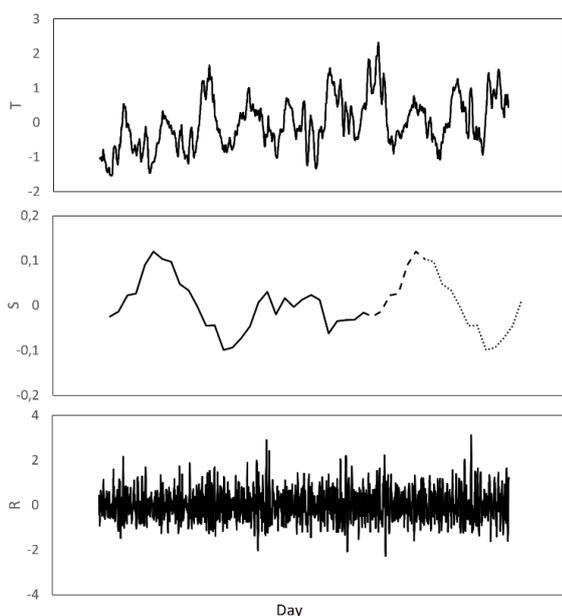


Рис. 1. Результаты аддитивной декомпозиции.

Анализ данных, представленных на рис.1 показывает, что трендовая составляющая *T<sub>c</sub>* течением времени, в целом, возрастает. Однако на ней можно выделить локальные максимумы и минимумы. Период локальных экстремумов примерно соответствует одному году. При этом, локальные максимумы примерно соответствуют летнему периоду. По-видимому, это связано с увеличением трафика в летний период.

Для сезонной составляющей *S* сплошная линия соответствует периоду сезонности  $p = 30$  суток, штриховая – периоду прогноза (7 суток), а пунктирная – повторению сезонной составляющей. Амплитуда изменения *S* почти на порядок меньше амплитуды случайной составляющей *R*, что подтверждает сделанный в [3] вывод о, преимущественно, случайном характере изменения среднего уровня аварийности.

Изменение ошибки обучения и валидации характеризуется быстрым снижением вплоть до 5 эпохи обучения независимо от составляющей временного ряда. Затем, до 10 эпохи обучения, включительно, происходит плавное снижение ошибок обучения и валидации практически до минимального значения. Более продолжительное обучение не приводит к заметному снижению ошибок, из чего можно сделать вывод о нецелесообразности обучения на протяжении более 10 эпох. Характер изменения разности между ошибками обучения и валидации соответствует отсутствию переобученности нейронных сетей. Уровни ошибок прогноза РНС2 [3] и РНС, входящих в состав каскадной нейросети приведен в табл. 1.

ТАБЛИЦА 1. ОШИБКА ПРОГНОЗА НЕЙРОННЫХ СЕТЕЙ ПОСЛЕ 10 ЭПОХ ОБУЧЕНИЯ

№	Нейронная сеть	Ошибка
1	РНС2[3]	0,6098
2	T	0,0994
3	S	0,0142
4	R	0,3921
5	Каскад	0,5057

Анализ приведенных в табл. 1 данных показывает, что применение каскада нейронных сетей позволяет снизить погрешность прогнозирования, по сравнению с РНС, обрабатывающей консолидированные данные, более чем на 0,1. При этом, минимальная погрешность наблюдается для сезонной составляющей, что объясняется малой вариативностью данных. Снижение ошибки прогнозирования сезонности до 0 возможно заменой использования РНС на выбор табличных значений ряда *S*. Несколько большая погрешность наблюдается для трендовой составляющей, что объясняется её тенденцией к непрерывному повышению и наличием локальных экстремумов. Одним из возможных путей снижения ошибки прогнозирования является выделение аналитически определенной линии тренда, например, с использованием регрессионного анализа, и периодической составляющей, отличающейся от сезонности периодом изменения 1 год и более, с последующим прогнозированием с помощью РНС периодической составляющей. Наибольшая ошибка наблюдается для прогноза случайной составляющей, однако, учет дополнительных факторов, таких как уровень солнечной активности и динамика уровня аварийности, позволит снизить ошибку прогноза *R*.

На рис. 2 представлены результаты прогнозирования составляющих временного ряда  $T$ ,  $S$  и  $R$ , и среднего уровня аварийности.

Анализ представленных на рис. 2 данных показывает, что прогнозные значения для трендовой  $T$  и сезонной  $S$  составляющей очень близки к истинным, а прогноз случайной составляющей  $R$ , несмотря на то что обеспечивает достаточно точное воспроизведение тенденции, по амплитуде изменения несколько отличается от реального. Тем не менее, суммарный прогноз среднего уровня аварийности у не просто воспроизводит тренд, наблюдающийся в реальности, но и оказывается достаточно близок к истинным значениям, особенно в начале интервала прогнозирования.

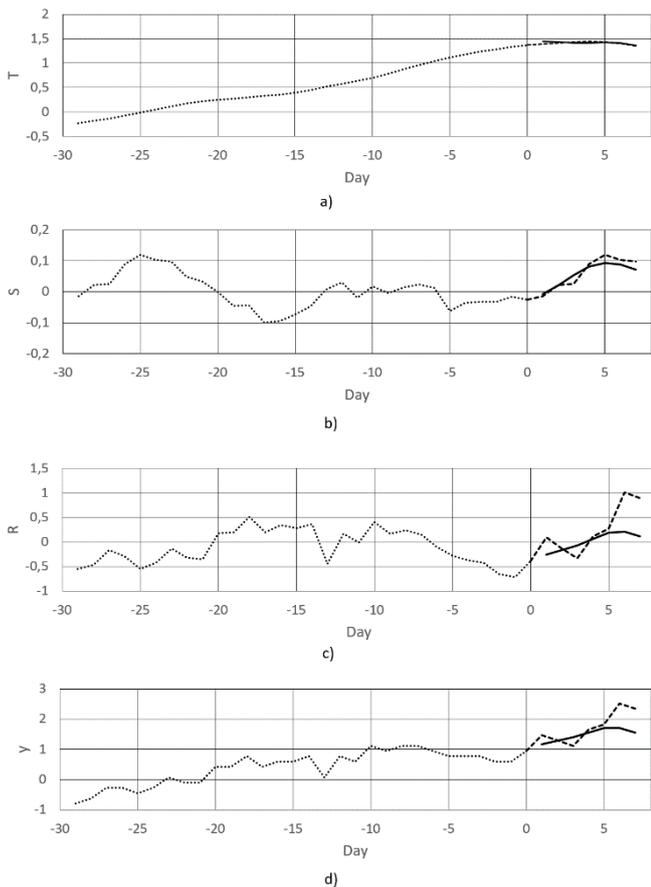


Рис. 2. Результаты прогнозирования составляющих: а) трендовой; б) сезонной; в) случайной д) среднего уровня аварийности; --- предшествующий период; - - - истинные значения; — — прогноз.

Также следует отметить, что замена значений сезонной составляющей, полученных с помощью РНС, на истинные значения, не приводит к существенному уточнению прогноза среднего уровня аварийности.

#### IV. ЗАКЛЮЧЕНИЕ

В целом, исходя из изложенного, можно сделать следующие выводы:

1. Аддитивная декомпозиция временного ряда показывает, преимущественно, случайный характер изменения среднего уровня аварийности на дорогах общего пользования.
2. Использование каскадной нейронной сети позволяет повысить точность прогноза среднего уровня аварийности более чем на 0,1. Ошибка прогноза при этом снижается с 0,6098 до 0,5057.
3. Замена прогнозирования сезонной составляющей  $S$  на её истинное значение приводит к заметному увеличению точности прогноза, однако позволяет снизить временные и вычислительные затраты на обучение и дальнейшее использование соответствующей РНС.
4. Незначительного увеличения точности прогнозирования можно добиться, разделив трендовую составляющую  $T$  на аналитически определенную линию тренда и периодическую составляющую.
5. Основные усилия по увеличению точности прогнозирования среднего уровня аварийности необходимо направить на учёт влияния дополнительных факторов на случайную составляющую  $R$ . К их числу относятся как факторы, влияющие на психофизиологическое состояние человека, так и динамика уровня аварийности.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] A.Mohammed, "Classification of Traffic Accident Prediction Models", A Review Paper, International Journal of Advances in Science Engineering and Technology, 2018, Vol 6, pp. 35-38.
- [2] A.Y. Abalyaev and L.V. Grunskaya, "Development of Recurrent Neuron Net System for Forecasting of an Accident Rate Mean Level on General Use Roads", 2022 International Conference on Engineering Management of Communication and Technology (EMCTECH), Vienna, Austria, 2022, pp. 1-4.
- [3] A.Y. Abalyaev and L.V. Grunskaya, "Neural Networks Features in the Task of Forecasting the Level of Accidents on the Public Roads," 2024 V International Conference on Neural Networks and Neurotechnologies (NeuroNT), Saint Petersburg, Russian Federation, 2024, pp. 91-92, doi: 10.1109/NeuroNT62606.2024.10585539.
- [4] Показатели состояния безопасности дорожного движения [Электронный ресурс], <http://stat.gibdd.ru> (03.04.2025 г.).
- [5] R.J. Hyndman and G. Athanasopoulos, "Forecasting: Principles and Practice", OTexts, 2021, 442 p.
- [6] M. Astrid Veronig, J. Shantanu, T. Podladchikova, W. Pötzi, and F. Clette, "Hemispheric Sunspot Numbers 1874-2020", Astronomy & Astrophysics, 2020, Vol. 652, pp. 56-57.