

DOI: 10.12731/2227-930X-2022-12-2-7-20**УДК 656.13****РОБАСТНОЕ ПРОГНОЗИРОВАНИЕ ИНТЕНСИВНОСТИ
ДВИЖЕНИЯ ТРАНСПОРТНЫХ ПОТОКОВ*****А.А. Власов***

Точная и своевременная информация о текущих и прогнозных параметрах транспортных потоков является важным условием функционирования интеллектуальных транспортных систем. Использование качественных данных позволит эффективнее решать задачу адаптивного управления транспортными потоками, снизить время движения, повысить точность планирования маршрутов поездок и в целом повысить эффективность использования транспортной инфраструктуры.

В данной статье представлено исследование воздействия светофорных объектов на регистрацию параметров транспортных потоков детекторами транспорта. Установлено, расположение зоны контроля детекторов транспорта в пределах участка формирования очереди транспортных средств ведет к искусственному зашумлению собираемых данных. Обосновано использование робастного прогноза параметров транспортного потока при решении задач адаптивного управления транспортными потоками.

Показано решение задачи робастного прогноза интенсивности движения с использованием модели LightGBM, основанной на алгоритме градиентного бустинга. Пространство признаков модели включало лаг интенсивности движения, дату и время полученных данных, признаки выходного дня и идентификатор группы полос движения. Робастность прогноза обеспечена за счет использования при вычислении функции ошибки модели сглаженного временного ряда методом LOWESS, при подаче на вход модели исходного, не сглаженного временного ряда.

Обученная на реальных данных модель имеет свойства робастности прогноза интенсивности движения, обеспечила фильтрацию стохастических колебаний и выбросов измеренных. При этом различает особенности суточных профилей для рабочих и выходных дней на всех группах полос, запаздывание прогноза наблюдалось только при наличии существенного отклонения наблюдаемой интенсивности от суточного тренда.

Ключевые слова: *робастный прогноз транспортного потока; градиентный бустинг; глубокое обучение; интеллектуальные транспортные системы; адаптивное управление транспортными потоками*

ROBUST FORECASTING OF TRAFFIC FLOW INTENSITY

A.A. Vlasov

Accurate and timely information about current and forecast parameters of traffic flows is an important condition for the functioning of intelligent transport systems. The use of high-quality data will make it possible to solve the problem of adaptive traffic flow management more effectively, reduce travel time, improve the accuracy of travel route planning, and generally improve the efficiency of using transport infrastructure.

This article presents studies of the impact of traffic light objects on the registration of traffic flow parameters by transport detectors. It is established that the location of the control zone of transport detectors within the area of the formation of the queue of vehicles leads to artificial noise of the collected data. The use of robust forecast of transport flow parameters in solving problems of adaptive traffic flow management is justified.

The solution of the problem of robust traffic intensity prediction using the LightGBM model based on the gradient boosting algorithm is shown. The feature space of the model included the traffic intensity log, the date and time of the data received, the weekend

features, and the lane group ID. The robustness of the forecast is ensured by using the smoothed time series model error function using the LOWESS method, when applying the original, non-smoothed time series to the model input.

The model trained on real data has properties of robustness of the traffic intensity forecast, provided filtering of stochastic fluctuations and outliers of the measured ones. At the same time, it distinguishes the features of the daily profiles for working and weekend days in all groups of bands, the forecast delay was observed only if there was a significant deviation of the observed intensity from the daily trend.

Keywords: *robust traffic flow prediction; gradient boosting; deep learning; intelligent transport systems; adaptive traffic flow control*

Введение

Снижение уровня загруженности дорог является приоритетной задачей для городов во всем мире, и в последние десятилетия существенное внимание уделено совершенствованию методов прогнозирования параметров транспортных потоков в контексте развития интеллектуальных транспортных систем (ИТС) и адаптивного управления транспортными потоками [1]. Эффективность ИТС в основном определяется качеством информации о дорожном движении, предоставляемой заинтересованным участникам дорожного движения, и способностью ее применять для разработки стратегии развития транспортной системы, систем управления и моделей дорожного движения.

Краткосрочное прогнозирование параметров транспортных потоков является важным элементом функционирования ИТС, при этом получение точного прогноза является сложной задачей, главным образом из-за динамического, сложного и стохастического характера движения транспорта. Сложности, связанные с прогнозированием транспортных потоков, определяются природой предметной области. Она включает множество ограничений, налагаемых транспортной инфраструктурой в виде пропускной

способности дорог, правил дорожного движения и управления движением, поведение отдельных агентов (участников дорожного движения), а также экзогенные факторы, такие как периоды активности участников движения (т. е. время суток, день недели и т.д.), погода, аварии и инциденты, закрытие дорог и т.д.

Набор данных для построения модели прогноза

Набор данных сформирован на основе измерений параметров транспортных потоков детекторами транспорта, установленными на пересечении проспекта Непокоренных и Гражданского проспекта г. Санкт-Петербург. Данные получены в период с марта по май 2020 года и включали интенсивность движения по полосам движения, агрегированные по 15 минут.

При решении задачи адаптивного управления светофорным объектом требуется дополнительная агрегация данных по группам полос (рисунок 1). Группы полос с номерами 3, 5 и 7 специализированы на пропуске левоповоротных потоков, а с номерами 1, 2, 4 и 6 – потоков прямого направления и совершающих правый поворот.

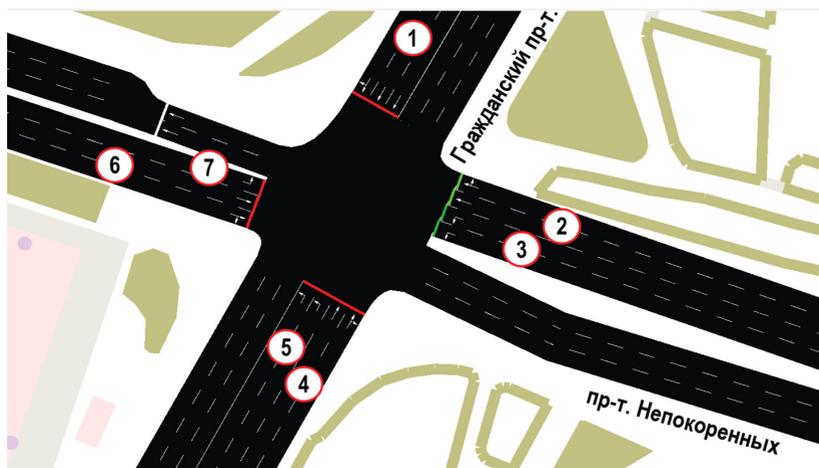


Рис. 1. Нумерация групп полос на пересечении проспекта Непокоренных и Гражданского проспекта г. Санкт-Петербург

Первичный анализ набора данных показал их существенную волатильность (рисунок 2). В относительном исчислении она составляет от 10 до 40%, что затрудняет использование полученных данных для принятия решений по управлению транспортным потоком без предварительной обработки.

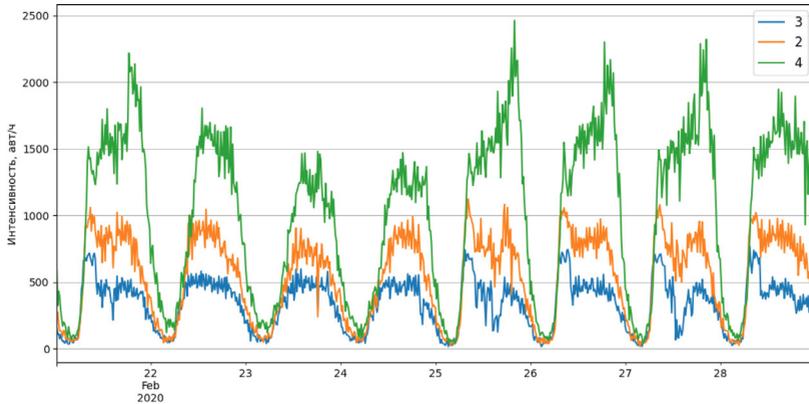


Рис. 2. Интенсивность движения по группам полос с 21.02.2020 по 29.02.2020

Набор данных содержат существенные выбросы в наблюдаемых данных, связанных по всей видимости, с работой детекторов транспорта (рисунок 3).

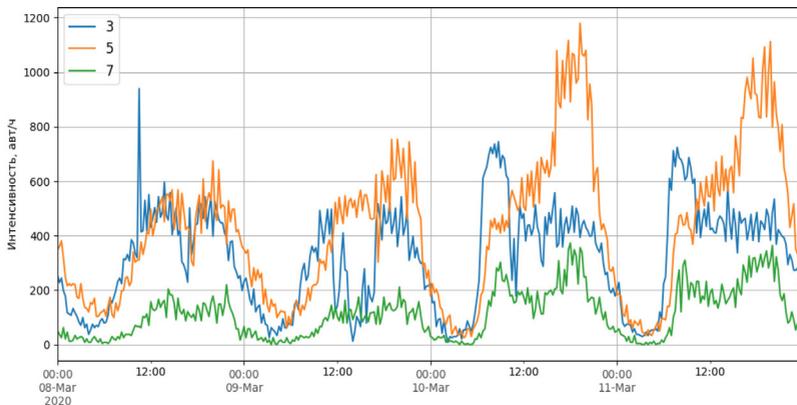


Рис. 3. Интенсивность движения по группам полос с 08.03.2020 по 12.03.2020

При построении моделей прогноза параметров транспортного потока исследователи стремятся наиболее точно воспроизвести наблюдаемые параметры, однако следует учитывать, что средства измерения могут вносить существенные искажения в наблюдаемый процесс.

Моделирование воздействия светофорных объектов на регистрацию интенсивность движения детекторами транспорта

Расположение детекторов транспорта на улично-дорожной сети, как правило, определяется решаемыми задачами по управлению транспортными потоками [7]. При адаптивном управлении светофорными объектами детекторы транспорта как правило располагаются на расстоянии от 60 до 180 м от пересечения. Детекторы транспорта фактически располагаются в зоне формирования очереди транспортных средств перед регулируемым перекрестком, что создает риск внесения существенных искажений в наблюдаемые параметры транспортного потока.

Для изучения влияния светофорных объектов на регистрируемую детекторами транспорта интенсивность движения использовано микроскопическое имитационное моделирование в программе SUMO (Simulation of Urban Mobility) [2]. Моделирование включало три этапа: подготовку транспортной сети, транспортного спроса и сценария моделирования; собственно, имитацию; составление отчетов и обработку результатов.

Для исследования использована модель пересечения ул. Измайлова и ул. Стрельбищенская г. Пензы (рисунок 4) со сгенерированным искусственно транспортным спросом. Интенсивность движения менялась каждый час моделирования по ул. Измайлова в диапазоне от 700 до 1500 авт/ч, по ул. Стрельбищенская от 300 до 600 авт/ч. Управление светофорным объектом производилось по алгоритму поиска разрыва в транспортном потоке.

Для регистрации параметров транспортного потока на расстоянии 50-60 м от стоп-линии были размещены виртуальные детек-

торы транспорта. По ул. Измайлова со стороны ул. Антонова на расстоянии 500 м от стоп-линии организована дополнительная зона детектирования. Агрегация данных производилась с периодом 300 с. Результаты имитации сохранялись в файл с расширением xml, который затем использован для анализа результатов.



Рис. 4. Микроскопическая модель пересечения ул. Измайлова и ул. Стрельбищевская г. Пензы

На рисунке 5 приведены результаты измерения интенсивности движения детекторами, расположенными на входе в перекресток.

Полученные значения интенсивность движения имеют существенные стохастические колебания, отклонение от истинных значений на измеренной интенсивности движения достигает 10-20%. Для зоны контроля, удаленной от светофорного объекта на 500 м. наблюдались выбросы измеренных значений на уровне 10-12% в моменты смены величины истинного транспортного спроса, в остальных периодах ошибка не превышала 2,5%.

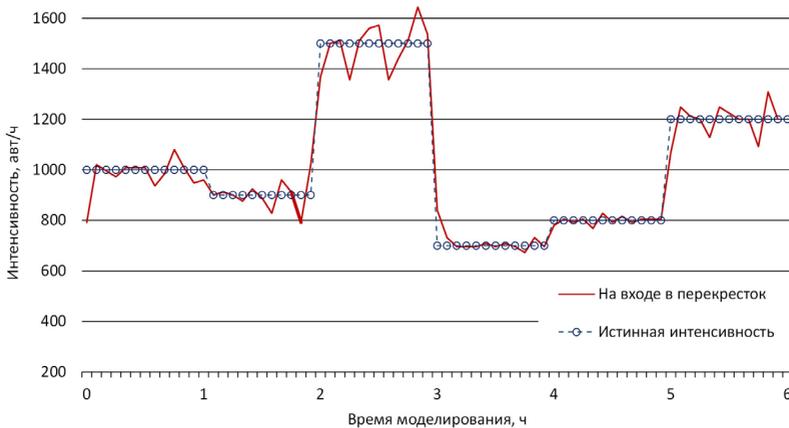


Рис. 5. Результаты измерения интенсивности движения на входе в перекресток

Размещение зоны учета параметров движения транспортных потоков в пределах участка формирования очереди транспортных средств ведет к зашумлению полученных данных, усилению стохастической составляющей. Результаты имитационного моделирования свидетельствуют о необходимости использования методов робастного прогноза параметров транспортного потока при решении задач адаптивного управления транспортной системой.

Методы получения краткосрочного прогноза параметров транспортного потока

Методы краткосрочного прогнозирования трафика в широком смысле подразделяются на параметрические и непараметрические подходы построения моделей прогноза [3, 4]. Наибольшую известность среди параметрических моделей имеет модель авторегрессионного интегрированного скользящего среднего (ARIMA) [4]. Однако основным предположением, используемым моделью ARIMA, является стационарность среднего, дисперсии и автокорреляции. Параметры транспортного потока, как правило, имеют выраженную

периодичность в суточном и недельном цикле, а также быстрые колебания, связанные с воздействием прочих внешних факторов. По этой причине ARIMA малоприменима для решения реальных задач прогнозирования параметров транспортного потока.

Непараметрические подходы предполагают использование методов машинного обучения для построения модели прогноза. Для решения задач прогноза параметров транспортного потока используются решения, основанные на ансамблях решающих деревьев (Random Forest) [5], градиентном бустинге [6], методе опорных векторов [7, 8] и нейронных сетях преимущественно с архитектурой LSTM или GRU [9].

В пространство признаков модели вводятся как структурированные переменные, такие как исторические данные интенсивности движения, погодные условия [10], так и слабо формализованные, как например информация из социальных сетей о наличии инцидентов [11, 12, 13, 14].

Обзор имеющихся решений не показал существенного превосходства ни одного из приведенных методов решения задачи прогнозирования параметров транспортных потоков. По всей видимости, определяющее значение имеет подготовка данных и набор переменных в пространстве признаков модели.

Разработка модели краткосрочного прогноза интенсивности движения

Для построения модели прогноза интенсивности движения использована библиотека LightAutoML [15], разработанная лабораторией искусственного интеллекта Сбербанка (Sberbank AI Lab AutoML). Данная библиотека построена на основе известных библиотек градиентного бустинга и линейной регрессии, имеет реализацию автоматической подготовки данных, настройки и оптимизация модели прогноза.

При проектировании модели стояла цель получения легко интерпретируемую структуру с достаточным для использования при адаптивном управлении транспортными потоками качеством про-

гноза. В пространство признаков модели были включены следующие переменные:

- текущая интенсивность движения, включающая историю из 10 отсчетов (лагов);
- дата и время получения данных. Кодирование данной переменной не производилось, что позволило LightAutoML сформировать дополнительные перепеменные модели, такие как признаки выходного дня.
- номер группы полос, так как интенсивность в течении для существенно различается по направлениям движения.

Обучение модели проводилось на 25 группах данных (*folds*), включающих обучающие и тестовые выборки. В качестве функции ошибки при обучении модели и метрики оценки результатов прогноза была выбрана MAE:

$$MAE = \frac{\sum_{i=1}^n |obs_i - pred_i|}{n},$$

где obs_i – наблюдаемая интенсивность движения на интервале i ; $pred_i$ – прогноз интенсивности движения на интервале i ; n – количество интервалов в выборке.

Для обеспечения робастности прогноза в качестве наблюдаемой интенсивности движения при вычислении функции ошибки модели использовался посуточно сглаженный ряд методом LOWESS [16]. При этом данные, подаваемые на вход модели прогноза брались из исходного, не сглаженного временного ряда.

Оптимизация с помощью LightAutoML структуры модели прогноза интенсивности движения дала блендинг двух методов LightGBM и LightGBM_tuned в соотношении 1:2.

Из первоначального набора данных выделены несколько последних дней, не используемых при обучении для которых выполнен контрольный прогноз интенсивности движения (рисунок 6).

Модель достаточно хорошо справилась с прогнозом интенсивности движения, различая особенности суточных профилей для рабочих и выходных дней на всех группах полос. Обеспечена ро-

бастность прогноза путем фильтрации стохастических колебаний и выбросов, измеренных детекторами транспорта данных. Запаздывание прогноза наблюдалось только при наличии существенно-го отклонения от суточного тренда.

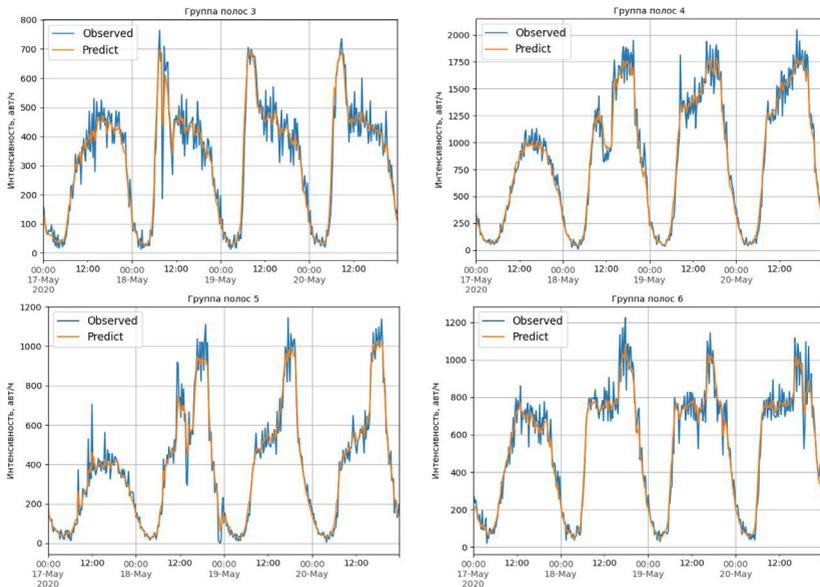


Рис. 6. Прогноз интенсивности движения по группам полос в период с 17.05.2020 по 21.05.2020: Observed – данные, полученные с детекторов транспорта; Predict – результаты прогноза интенсивности движения

Заключение

Краткосрочный прогноз параметров транспортных потоков, включая интенсивность движения является сложной технической задачей в следствие сложности транспортной системы как объекта наблюдения и управления в целом. Кроме того, данные о параметрах транспортных потоков в городах, содержат существенный шум, привнесенный особенностями размещения и функционирования системы контроля параметров транспортных потоков.

Модель краткосрочного прогноза параметров транспортного потока должна иметь функции робастности и обеспечивать фильтрацию

шумов искусственного происхождения и выбросов, возникающих в результате сбоев в работе детекторов транспорта. Указанным требованиям отвечают модели, основанные на градиентном бустинге и обученные с использованием предложенного подхода к формированию признаков модели и вычислению функции ошибки.

Полная оценка качества прогноза интенсивности движения требует проведения дополнительных исследований, включающих имитацию работы систем адаптивного управления транспортными потоками и оценку задержки транспортных средств.

Благодарности. Выражаем благодарность ООО «РИПАС» и лично Подозерову Н.Е. за предоставленные данные интенсивности движения.

Список литературы / References

1. Essien, A., Petrounias, I., Sampaio, P., Sampaio, S.: Deep-PRESIMM: Integrating Deep Learning with Microsimulation for Traffic Prediction. *IEEE International Conference on Systems, Man, and Cybernetics. IEEE Xplore*, 2019, pp. 1–6.
2. Krajzewicz, D., Erdmann, J., Behrisch, M., Bieker, L.: Recent Development and Applications of SUMO – Simulation of Urban Mobility. *International Journal on Advances in Systems and Measurements*, 2012, vol. 5 (3&4), pp.128-138.
3. Lana, I., Del-Ser, J., Velez, M., Vlahogianni, E.I.: Road traffic forecasting: recent advances and new challenges. *IEEE Intell. Transp. Syst. Mag.*, 2018, vol. 10, pp. 93–109. <https://doi.org/10.1109/MITS.2018.2806634>
4. Smith, B.L., Williams, B.M., Keith Oswald, R.: Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 2002, vol. 10, pp. 303–321. [https://doi.org/10.1016/S0968-090X\(02\)00009-8](https://doi.org/10.1016/S0968-090X(02)00009-8)
5. Liu, Y., Wu, H.: Prediction of Road Traffic Congestion Based on Random Forest. *2017 10th International Symposium on Computational Intelligence and Design (ISCID)*. IEEE, 2017, pp. 361–364.

6. Dong, X., Lei, T., Jin, S., Hou, Z.: Short-Term Traffic Flow Prediction Based on XGBoost. *2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS)*. IEEE, 2018, pp. 854–859.
7. Wang, X., An, K., Tang, L., Chen, X.: Short term prediction of free-way exiting volume based on SVM and KNN. *International Journal of Transportation Science and Technology*. 2015. <https://doi.org/10.1260/2046-0430.4.3.337>
8. Philip, A.M., Ramadurai, G., Vanajakshi, L.: Urban arterial travel time prediction using support vector regression. *Transportation in Developing Economies*, 2018, vol. 4, 7. <https://doi.org/10.1007/s40890-018-0060-6>
9. Fu R., Zhang Z., Li L. Using LSTM and GRU neural network methods for traffic flow prediction. Chinese Association of Automation, 2017, pp. 324-328.
10. Essien, A., Petrounias, I., Sampaio, P., Sampaio, S.: The impact of rainfall and temperature on peak and off-peak urban traffic. *International Conference on Database and Expert Systems Applications*. Springer, Cham. 2018, pp. 399–407.
11. Goh, G., Koh, J., Zhang, Y.: Twitter-Informed Crowd Flow Prediction. *2018 IEEE Conference on Data Mining Workshops (ICDMW)*. 2018. pp. 624–631.
12. Ni, M., He, Q., Gao, J.: Using social media to predict traffic flow under special event conditions. *The 93rd annual meeting of transportation research board*. 2014.
13. Wongcharoen, S., Senivongse, T.: Twitter analysis of road traffic congestion severity estimation. *2016 13th International Joint Conference on Computer Science and Software Engineering, JCSSE 2016*. IEEE 2016, pp. 1–6.
14. Essien, A., Petrounias, I., Sampaio, P. et al. A deep-learning model for urban traffic flow prediction with traffic events mined from twitter. *World Wide Web*. 2020. <https://doi.org/10.1007/s11280-020-00800-3>
15. LightAutoML – automatic model creation framework. URL: www.github.com/sberbank-ai-lab/LightAutoML (дата обращения 30.05.2021)

16. Cleveland, William S.. LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *The American Statistician*. 1981, vol. 35 (1), 54. <https://doi.org/10.2307/2683591>

ДАнные ОБ АВТОРЕ

Власов Алексей Александрович, доцент кафедры «Организация безопасности движения», кандидат технических наук
ФГБОУ ВО Пензенский государственный университет архитектуры и строительства
ул. Германа Титова, 28, г. Пенза, Пензенская область, 440028, Российская Федерация
vlasov_a71@mail.ru

DATA ABOUT THE AUTHOR

Alexey A. Vlasov, Associate Professor «Organization of traffic safety»,
Candidate of Technical Sciences
Penza State University of Architecture and Building
28, Herman Titov Str., Penza, Penza Region, 440028, Russian Federation
vlasov_a71@mail.ru

Поступила 11.03.2022

После рецензирования 25.03.2022

Принята 07.04.2022

Received 11.03.2022

Revised 25.03.2022

Accepted 07.04.2022