

УДК 656.021, 656.11

Гантимурова Юлия Олеговна,  
к.т.н., доцент, доцент кафедры «Управление на автомобильном транспорте»,  
ФГБОУ ВО «Ангарский государственный технический университет»,  
e-mail: juliapoltavskaya@mail.ru

## МОДЕЛЬ ПРОГНОЗИРОВАНИЯ ТРАНСПОРТНЫХ ЗАТОРОВ НА ОСНОВЕ GPS-ДААННЫХ

Gantimurova Y.O.

### GPS-BASED TRAFFIC CONGESTION PREDICTION MODEL

**Аннотация.** В статье представлен комплекс методов и моделей для прогнозирования транспортных заторов, основанный на анализе GPS-треков и данных о дорожной сети. Рассмотрены этапы предварительной обработки данных, построения моделей оценки скорости и классификации уровня загруженности. Предлагается система автоматического обновления информации в реальном времени, что позволяет своевременно реагировать на изменения дорожной ситуации, снижая заторы и повышая безопасность движения.

**Ключевые слова:** транспортный затор, GPS треки, дорожное движение, модель прогнозирования, скорость движения, машинное обучение.

**Abstract.** This article presents a set of methods and models for predicting traffic congestion based on the analysis of GPS tracks and road network data. It examines the stages of data preprocessing, the construction of speed estimation models, and congestion classification. A system for automatically updating information in real time is proposed, enabling timely responses to changing traffic conditions, reducing congestion and improving traffic safety.

**Keywords:** traffic congestion, GPS tracks, road traffic, pattern prediction, traffic speed, machine learning.

Транспортные заторы – одна из ключевых проблем современных мегаполисов, требующая поиска эффективных решений [1, 2]. Основные причины возникновения заторов можно условно разделить на две категории. Первая связана с превышением пропускной способности дорожной сети – когда количество транспортных средств на дороге превышает её возможности, что приводит к образованию узких мест и заторам. Вторая – дорожно-транспортные происшествия и неблагоприятные погодные условия, которые замедляют движение и вызывают скопление транспорта. Длительные задержки в движении создают неудобства для людей и ухудшают экологическую ситуацию из-за роста выбросов углекислого газа [1].

В связи с увеличением числа транспортных средств, строительство новых дорог и автомагистралей не всегда позволяет полностью решить проблему, поскольку такие меры требуют значительных временных и финансовых затрат. Поэтому актуальной становится разработка интеллектуальных систем и моделей для прогнозирования транспортных заторов и оптимизации дорожного движения [2, 3-5].

Анализ дорожного движения можно проводить на различных уровнях детализации: микроскопическом, мезоскопическом и макроскопическом [6]. Микроскопический анализ фокусируется на поведении отдельных транспортных средств, мезоскопический – на отдельных участках и перекрестках, а макроскопический – на транспортной системе города или региона. В современных исследованиях особое внимание уделяется сбору и обработке данных о движении транспортных средств для построения моделей прогнозирования.

Для сбора данных о состоянии транспортного потока используются различные методы. Традиционные системы включают пневматические датчики, индукционные петли, радар-детекторы и видеонаблюдение, однако они требуют значительных инвестиций и технического обслуживания. Альтернативой является использование GPS-датчиков, установленных внутри транспортных средств, что позволяет получать данные о скорости, времени и маршрутах движения в реальном времени. Такой подход обеспечивает более гибкое и масштабируемое решение для мониторинга дорожной ситуации [7, 8].

Географические данные, полученные с помощью GPS-устройств, широко используются для анализа перемещений населения и транспортных средств. GPX (GPS Exchange Format) — популярный формат для хранения таких данных. Прогнозирование движения на основе треков позволяет улучшить навигационные системы, оптимизировать транспортные маршруты и повысить безопасность. Обработка собранных данных включает их проверку и структурирование. В частности, из GPS-траекторий выделяют важные параметры, такие как скорость движения на различных участках дороги, а также время, расстояние и координаты.

Рассмотрим алгоритм предварительной обработки данных с GPS-устройств, который включает следующие шаги.

1. Очистка данных: удаление выбросов и аномальных точек, таких как выпадающие или ошибочные координаты, вызванные сбоями в работе устройств или помехами сигнала.

2. Обработка пропущенных данных: заполнение пропусков или удаление участков с недостаточной точностью, чтобы обеспечить целостность набора данных.

3. Уменьшение размерности: применение методов снижения сложности данных, таких как фильтрация или кластеризация, для выделения ключевых характеристик траектории.

4. Фильтрация объектов: исключение из анализа нежелательных элементов, например, остановок вне маршрута или случайных точек, не относящихся к движению транспортных средств.

5. Структурирование: организация данных в удобные для анализа форматы, такие как таблицы или временные ряды, с учетом маршрутов, времени и других параметров.

Эти этапы позволяют получить упорядоченный набор данных, который может быть использован для дальнейшего анализа, моделирования и прогнозирования транспортных процессов. На рисунке 1 приведен пример трека маршрута на картах Google Earth. Информация сохраняется с интервалом в одну секунду и содержит следующие данные: время, высота, координаты (широта и долгота) и скорость. На рисунке 2 – структурированный набор данных в MS Excel после извлечения и обработки.



Рисунок 1 – Трек маршрута в программе Google Earth с отображением основных характеристик движения

	A	B	C	D	E	F	G	H	I
	номер трека	количество точек записи	высота	длина отрезка	время отрезка	скорость на отрезке	направление	дата/время	позиция
1	1 (1160)	1	479 м	18 м	0:00:01	66,00 334.8° истина		26.04.2023 13:32	N52° 21.532' E104° 07.776'
2		2	479 м	19 м	0:00:01	68,00 334.1° истина		26.04.2023 13:32	N52° 21.541' E104° 07.769'
3		3	479 м	18 м	0:00:01	65,00 332.8° истина		26.04.2023 13:32	N52° 21.550' E104° 07.762'
4		4	479 м	19 м	0:00:01	67,00 330.0° истина		26.04.2023 13:32	N52° 21.559' E104° 07.755'
5		5	479 м	18 м	0:00:01	66,00 325.7° истина		26.04.2023 13:32	N52° 21.568' E104° 07.746'
6		6	478 м	19 м	0:00:01	68,00 328.6° истина		26.04.2023 13:32	N52° 21.576' E104° 07.737'
7		7	478 м	19 м	0:00:01	67,00 327.9° истина		26.04.2023 13:32	N52° 21.585' E104° 07.728'
8		8	478 м	19 м	0:00:01	70,00 326.0° истина		26.04.2023 13:32	N52° 21.593' E104° 07.720'
9		9	477 м	19 м	0:00:01	69,00 323.2° истина		26.04.2023 13:32	N52° 21.602' E104° 07.710'
10		10	477 м	19 м	0:00:01	70,00 324.2° истина		26.04.2023 13:32	N52° 21.610' E104° 07.700'
11		11	476 м	19 м	0:00:01	68,00 321.2° истина		26.04.2023 13:32	N52° 21.619' E104° 07.690'
12		12	476 м	19 м	0:00:01	69,00 321.3° истина		26.04.2023 13:32	N52° 21.627' E104° 07.679'
13		13	476 м	18 м	0:00:01	66,00 321.5° истина		26.04.2023 13:32	N52° 21.635' E104° 07.669'
14		14	475 м	19 м	0:00:01	69,00 319.3° истина		26.04.2023 13:32	N52° 21.642' E104° 07.659'
15		15	475 м	20 м	0:00:01	71,00 313.2° истина		26.04.2023 13:32	N52° 21.650' E104° 07.648'
16		16	474 м	18 м	0:00:01	65,00 311.3° истина		26.04.2023 13:32	N52° 21.657' E104° 07.635'
17		17	473 м	19 м	0:00:01	68,00 312.2° истина		26.04.2023 13:32	N52° 21.664' E104° 07.623'
18		18	472 м	19 м	0:00:01	69,00 300.6° истина		26.04.2023 13:32	N52° 21.671' E104° 07.611'
19		19	471 м	19 м	0:00:01	69,00 300.6° истина		26.04.2023 13:32	N52° 21.676' E104° 07.596'
20		20	470 м	20 м	0:00:01	73,00 296.1° истина		26.04.2023 13:33	N52° 21.681' E104° 07.581'
21		21	469 м	21 м	0:00:01	76,00 294.9° истина		26.04.2023 13:33	N52° 21.686' E104° 07.565'
22		22	468 м	21 м	0:00:01	77,00 293.2° истина		26.04.2023 13:33	N52° 21.691' E104° 07.549'
23		23	467 м	22 м	0:00:01	80,00 294.8° истина		26.04.2023 13:33	N52° 21.695' E104° 07.531'
24		24	466 м	23 м	0:00:01	83,00 295.3° истина		26.04.2023 13:33	N52° 21.701' E104° 07.513'
25		25	465 м	23 м	0:00:01	82,00 294.3° истина		26.04.2023 13:33	N52° 21.706' E104° 07.495'
26		26	465 м	23 м	0:00:01	82,00 291.7° истина		26.04.2023 13:33	N52° 21.711' E104° 07.477'
27		27	464 м	23 м	0:00:01	83,00 288.4° истина		26.04.2023 13:33	N52° 21.715' E104° 07.458'
28		28	463 м	23 м	0:00:01	83,00 288.4° истина		26.04.2023 13:33	N52° 21.719' E104° 07.439'
29		29	463 м	22 м	0:00:01	78,00 285.9° истина		26.04.2023 13:33	N52° 21.723' E104° 07.419'
30		30	462 м	22 м	0:00:01	79,00 284.6° истина		26.04.2023 13:33	N52° 21.726' E104° 07.401'
31		31	462 м	24 м	0:00:01	85,00 278.2° истина		26.04.2023 13:33	N52° 21.729' E104° 07.382'
32		32	461 м	22 м	0:00:01	81,00 274.6° истина		26.04.2023 13:33	N52° 21.731' E104° 07.362'
33		33	460 м	23 м	0:00:01	82,00 273.1° истина		26.04.2023 13:33	N52° 21.732' E104° 07.342'

Рисунок 2 – Структурированный набор данных о характеристиках движения по маршруту в MS Excel

Предварительно обработанная траектория может быть выражена следующим образом:

$$T_i \in \{t_i, lat_i, lon_i, dir_i, v_i\}, \quad (1)$$

где  $T_i$  – конкретная траектория движения автомобиля,  $t_i$  – время суток в местном часовом поясе,  $lat_i, lon_i$  – информация о местоположении (широта, долгота),  $dir_i$  – направление движения,  $v_i$  – текущая скорость автомобиля.

После предварительной обработки GPS треков набор данных готов к анализу.

Карта является набором данных, представляющих детализацию конкретного местоположения с подробным описанием основных дорог и инфраструктурных объектов. Данные карты можно разделить на более мелкие наборы данных. Предположим, что дорога  $R$  представляет собой набор сегментов дороги  $e$ :

$$R \in \{e_1, e_2, \dots, e_n\}, \quad (2)$$

где каждый сегмент дороги состоит из начальной и конечной точек, и в свою очередь:

$$e \in \{lat_1, lon_1, lat_2, lon_2\}, \quad (3)$$

где  $lat_1, lon_1$  представляет начальную точку, а  $lat_2, lon_2$  – конечную точку сегмента дороги [5].

Рассмотрим модель прогнозирования дорожных заторов, в основу которой заложены данные GPS треков. Она включает два основных компонента: модель оценки скорости движения и модель прогнозирования дорожных заторов (рисунок 3). Процесс начинается с предварительной обработки данных GPS, собранных с движущихся транспортных средств. Помимо этого используется информация о дорожной сети, включая расположение регулируемых пересечений, остановочных пунктов на маршруте, которая служит дополнительным источником данных для модели оценки скорости [5, 9].

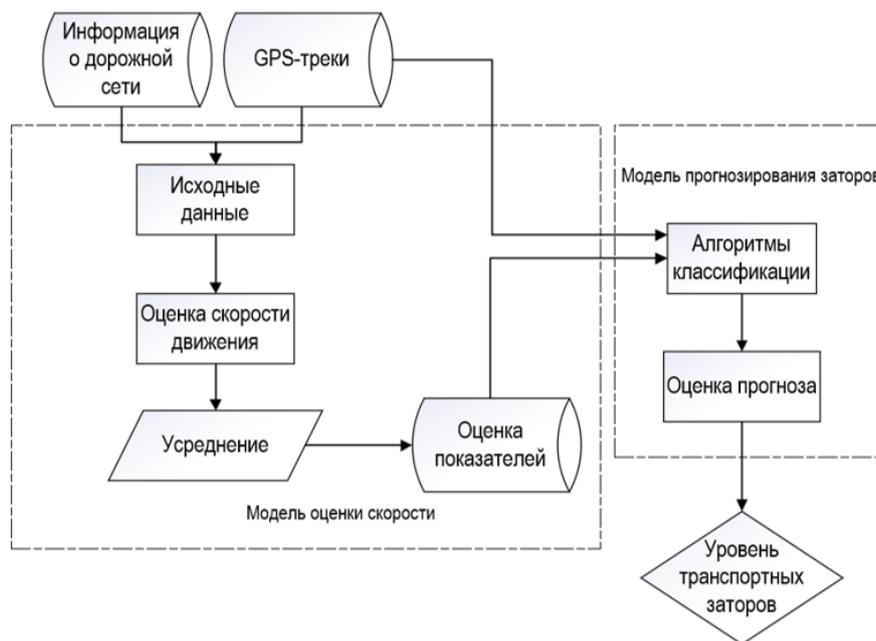


Рисунок 3 – Модель прогнозирования транспортных заторов

Расчетная скорость движения для каждого участка дороги сохраняется в специальной матрице, которая используется для дальнейшей обработки и анализа. Входные данные – оперативные GPS-данные и результаты модели оценки скорости подаются в алгоритмы классификации, предназначенные для прогнозирования уровня загруженности дорожных участков.

Для оценки точности и надежности модели применяется схема перекрестной проверки K-Fold, которая помогает избежать переобучения и обеспечивает более объективную оценку производительности модели. Итоговые результаты прогнозирования отображаются на цифровой карте, что позволяет участникам дорожного движения своевременно получать информацию о текущей ситуации и уровне загруженности дорог.

Дополнительно в системе реализованы механизмы автоматического обновления данных в реальном времени, что обеспечивает актуальность информации и помогает в управлении дорожным движением, снижая заторы и повышая безопасность на дорогах. Для анализа траекторий движения транспортных средств анализируемые маршруты разделяются на сегменты в соответствии с координатами по Гринвичу, которые затем преобразуются в декартовы координаты. Каждому сегменту подбирается полиномиальная функция методом наименьших квадратов, для моделирования движения по участку [5, 7, 10, 11].

Из-за снижения интенсивности движения в определённые периоды суток, возникают участки дороги, по которым транспортные средства не проходят в этот промежуток времени. Для решения этой проблемы используется наложение исторических траекторий GPS, что позволяет определить среднюю скорость движения на каждом участке дороги. Статистические показатели, такие как средняя скорость, стандартное отклонение, минимальная и максимальная скорость, рассчитываются каждые 5 минут и сохраняются для анализа. На основе статистических данных, текущей информации и коэффициентов задержки движения на каждом участке, скорость движения регулярно рассчитывается и обновляется. Это обеспечивает актуальную оценку скорости и помогает в дальнейшем прогнозировании дорожных условий.

После того, как средняя скорость движения по всему маршруту сохранена в матрице по сегментам, прогнозирование заторов завершено. Исходные данные GPS упрощаются с помощью предварительного анализа (п. 3 алгоритма), и рассчитывается предполагаемая скорость движения по маршруту.

Прогнозирование заторов осуществляется с использованием древовидных классификаторов, а именно алгоритма дерева принятия решений. Данные GPS-треков из модели оценки скорости движения используются в качестве обучающих и тестовых данных для подгонки моделей. Обучающие данные включают информацию о времени, скорости,

направлении, пробеге и местоположении (широта, долгота), а также об уровне заторов. Поскольку связь между независимыми переменными и целевой переменной имеет ключевое значение для классификации, выполняется анализ обучающих данных. Это необходимо для того, чтобы удостовериться, что при использовании моделей машинного обучения учитываются только те признаки, которые действительно важны.

Реализованные модели позволяют

своевременно оценивать текущую ситуацию на дорогах и предсказывать уровень загруженности участков в реальном времени. Такой подход способствует оптимизации управления дорожным движением, снижению заторов и повышению безопасности на дорогах. Внедрение автоматизированных систем на базе предложенных методов может значительно улучшить качество транспортных потоков и повысить комфорт для участников дорожного движения [1, 4].

## СПИСОК ЛИТЕРАТУРЫ

1. **Лебедева, О.А.** Транспортная инфраструктура как основополагающий фактор эффективного функционирования экономики страны / О. А. Лебедева, Ю. О. Полтавская, З. Н. Гаммаева, Т. В. Кондратенко. – Текст: непосредственный // Сборник научных трудов Ангарского государственного технического университета. – 2018. – Т. 1. № 15. – С. 125-130.
2. **Лагерева, Р.Ю.** Методика предупреждения сетевых транспортных заторов / Р. Ю. Лагерева, А. Ю. Михайлов, С. В. Лагерева. – Текст: непосредственный // Вестник НЦБЖД. – 2010. – № 5. – С. 82-88.
3. **Полтавская, Ю.О.** Оценка условий движения транспортных потоков с применением геоинформационных технологий / Ю. О. Полтавская, М. Н. Крипак, В. Е. Гозбенко. – Текст: непосредственный // Современные технологии. Системный анализ. Моделирование. – 2016. – № 1 (49). – С. 155-161.
4. **Полтавская, Ю.О.** Применение геоинформационных систем для обеспечения устойчивого развития транспортной системы города / Ю. О. Полтавская. – Текст: непосредственный // В сборнике: Информационные технологии в науке, управлении, социальной сфере и медицине. Сборник научных трудов VI Международной научной конференции. Под редакцией О.Г. Берестневой, В.В. Спицына, А.И. Труфанов, Т.А. Гладковой. – 2019. – С. 164-167.
5. **Kyaw, T.** Predicting on-road traffic congestion from public transport GPS data / T. Kyaw, N. N. Oo, W. Zaw // International Journal of Advances in Scientific Research and Engineering. – Vol 6 (3). – 2020. – p. 233-241.
6. **Mahmud, K.** Possible causes and solutions of traffic jam and their impact on the economy of Dhaka city / K. Mahmud, K. Gope, S. M. R. Chowdhury // Journal of Management and Sustainability. – Vol. 2(2). – 2012. – 112 p.
7. **Пиров, Ж.Т.** Влияние распределения транспортных потоков на скорость сообщения на сегментах городских улиц с регулируемым движением / Ж. Т. Пиров, А. Ю. Михайлов. – Текст: непосредственный // Интеллект. Инновации. Инвестиции. – 2020. – № 2. – С. 115-124.
8. **Kerner, B.S.** Introduction to modern traffic flow theory and control / B. S. Kerner // Springer-Verlag Berlin Heidelberg. – 2009. – p 45.
9. **Полтавская, Ю.О.** Моделирование продолжительности движения по маршруту с учетом характеристик улично-дорожной сети / Ю. О. Полтавская, О. А. Лебедева. – Текст: непосредственный // В книге: Новые информационные технологии в исследовании сложных структур. материалы Тринадцатой Международной конференции. Томский государственный университет. Томск. – 2020. – С. 101-102.
10. **Михайлов, А.Ю.** Интегральный критерий оценки качества функционирования улично-дорожных сетей / А. Ю. Михайлов. – Текст: непосредственный // Известия Иркутской государственной экономической академии. – 2004. – № 2. – С. 50-53.
11. **Крипак, М.Н.** Оценка состояния улично-дорожной сети крупного города / М. Н. Крипак, О. А. Лебедева. – Текст: непосредственный // Современные технологии. Системный анализ. Моделирование. – 2016. – № 3 (51). – С. 171-174.