

УДК 004.942:656.021.2

МОДЕЛИРОВАНИЕ ЛОКАЛЬНЫХ ТРАНСПОРТНЫХ ПОТОКОВ В УСЛОВИЯХ НЕПОЛНОТЫ ИСХОДНЫХ ДАННЫХ

МИТРОХИН Максим Александрович, д-р техн. наук, профессор; e-mail: mmax83@mail.ru

АЛЯЕВ Александр Олегович, лаборант-исследователь; e-mail: alyaev-2003@mail.ru

ЛОБАНОВ Рамиль Ильясович, студент; e-mail: ramil_lobanov@mail.ru

КИРЬЯНОВ Владислав Евгеньевич, студент; e-mail: vlad.kiryanov.2019@mail.ru

Пензенский государственный университет, Кафедра «Вычислительная техника», Пенза

Целью настоящей статьи является проверка гипотезы о возможности эффективного моделирования дорожного трафика с использованием ограниченных наборов данных. Конкретно: исследование фокусируется на возможности моделирования потоков на перекрестках, основываясь исключительно на информации о движении транспорта в смежных участках дорожной инфраструктуры.

Одним из главных этапов разработки системы было исследование и выбор оптимального интервала усреднения данных о количестве транспортных средств, движущихся в модели. Этот выбор является основным, так как он непосредственно влияет на точность и эффективность работы системы. В процессе анализа различных вариантов интервалов усреднения была выявлена оптимальная конфигурация, которая обеспечивает наилучший баланс между точностью данных и эффективностью системы. На основе результатов моделирования был выбран оптимальный интервал усреднения данных.

В рамках этой цели основными задачами в данной работе являются: разработка программы автоматического запуска процесса моделирования при поступлении новых данных с камеры, проверка адекватности моделирования на примере одного конкретного перекрестка, а также проверка гипотезы относительно возможности эффективного моделирования перекрестков при использовании ограниченных объемов данных.

Разработанная система моделирования транспортных потоков представляет собой решение, которое базируется на использовании данных, получаемых с интеллектуальной камеры, для анализа и оптимизации движения транспортных средств. Эта система была создана в рамках исследовательского проекта, который имел своей основной целью оптимизацию процесса анализа транспортного потока в городских и дорожных условиях.

Разработанная модель обладает широким спектром потенциальных применений. Она может быть интегрирована в городское планирование с целью создания эффективных моделей дорожного движения, которые впоследствии могут быть использованы для эффективного управления дорожным движением и оптимизации дорожной инфраструктуры.

Ключевые слова: моделирование транспортных потоков; трафик; симулятор городской мобильности; интервал усреднения данных.

DOI: 10.20295/2412-9186-2023-9-04-355-367

▼ Введение

Рост городского населения и увеличение автотранспортного движения приводят к необходимости эффективного управления транспортной инфраструктурой и оптимизации городского движения. Важность эффективного управления транспортом становится все более очевидной, так как увеличение трафика может приводить к дорожным заторам, которые, в свою очередь, приводят к ухудшению экологической обстановки и экономическому ущербу из-за потери времени, опозданий.

Для выявления рациональных схем управления движением транспорта, анализа причин появления дорожных заторов и способов их

устранения широко применяется имитационное моделирование [1].

В исследованиях отечественных [2] и иностранных авторов [3] обозначена перспективность имитационного моделирования при работе с транспортными системами.

В работе [4] авторы отмечают высокую точность различных алгоритмов имитационного моделирования. В [5] отмечено, что модели симуляции дорожного движения могут стать отличным вариантом для оценки управления городским движением.

В исследовании Т. А. Кузнецова [6] по транспортному моделированию рассмотрена задача оптимизации времени действия зеленого

сигнала на светофорах у перекрестков. Используя имитационное моделирование, удалось оптимизировать параметры работы светофора, сократив время пребывания транспорта на перекрестке и уменьшив объем заторов.

В рамках изучения современных инструментов имитационного моделирования транспортных потоков авторы Ю. И. Захарова и Е. С. Карнауха [7] подтвердили значительное упрощение процесса изучения и контроля динамики транспортного потока средствами имитационного моделирования.

Немаловажным аспектом имитационного моделирования является его точность. В [8] предлагается инновационная структура калибровки имитационных моделей, обеспечивающая одновременную настройку эксплуатационных параметров и параметров безопасности, улучшая производительность модели и позволяя оценивать реальные дорожные конфликты.

В настоящее время на рынке представлены несколько программных продуктов, средствами которых можно моделировать трафик. Это *SUMO (Simulation of Urban Mobility)*, *Avenue 2.0*, *LISA+*, *INES+* и другие. Наиболее широкие возможности для моделирования трафика и интеграции с другими программными продуктами имеет *SUMO*. К достоинствам данной программы, по отношению к вышперечисленным аналогам, можно отнести ее гибкость настроек и возможностей моделирования, реалистичность воспроизведения, а также распространение на бесплатной основе.

Одним из основных этапов моделирования является задание информации о параметрах трафика, в частности об интенсивности, если речь идет о воспроизведении функционирования реального участка дорожной сети. Информацию о трафике можно получать различными способами, такими как: датчики на дорогах, камеры наблюдения, данные GPS, индуктивные датчики, реагирующие на металлические объекты, такие как машины.

Рассматривая тему анализа технологий сбора данных о параметрах движения транспортных потоков в рамках научной работы, А. А. Ромашова [9] представила существенный анализ современных методов сбора информации о движении автотранспорта, приведя

достоинства и недостатки каждого метода подсчета.

Наиболее перспективным вариантом является использование интеллектуальных видеокамер, способных вести подсчет проезжающих транспортных средств по каждому из направлений движения. Один из вариантов такой камеры¹ работает в составе информационной системы² сбора и анализа данных о транспортных потоках. Такая система позволяет распознавать 13 типов транспортных средств, определять их скорость и направление движения, рассчитывать интенсивность транспортного потока в реальном времени. Данные системы предоставляются в формате: тип транспортного средства (ТС); позиция старта; позиция финиша; дата и время записи информации о проезде. Аналогичная система³ требует минимального вмешательства человека в процесс, а также способна распознавать транспортные средства на полосах движения. Существенным достоинством использования интеллектуальных камер является отсутствие необходимости проведения монтажных работ под дорожным полотном, как в случае с индуктивными датчиками, или установки датчиков непосредственно на транспортных средствах. Однако, несмотря на достоинства интеллектуальных видеокамер, в некоторых случаях возможность их установки может быть ограничена отсутствием подходящих инженерных конструкций, ограниченным обзором, невозможностью подведения электропитания или высокоскоростного интернета и т. д. В этом случае возникает ситуация неполноты исходных данных о транспортных потоках, необходимых для имитационного моделирования. Эта ситуация заключается в том, что для части участков дорожно-транспортной сети информация о транспортных потоках отсутствует ввиду невозможности установки средств сбора информации или сокращения затрат на установку в условиях ограниченных

¹Расчет транспортного потока на основе YOLOv5 и DeepSORT на базе Deepstream. — URL <https://habr.com/ru/articles/725916/>.

²SMART TRAFFIC SYSTEM. — URL <http://traffic.codeinside.ru/>.

³Интеллектуальная транспортная система. — URL: <https://neurocore.ru/blogs/traffic>.

муниципальных бюджетов. Особенно чувствительной к неполноте исходных данных является задача локального управления, например оптимизации светофорного регулирования на основе имитационного моделирования.

Цель данной статьи заключается в оценке адекватности имитационной модели транспортных потоков в условиях неполноты исходных данных. Более конкретно: исследование фокусируется на возможности создания имитационной модели транспортных потоков на перекрестках, используя информацию о движении транспорта в примыкающих элементах дорожной инфраструктуры. В качестве критерия адекватности предлагается использовать среднеквадратичную ошибку расхождения данных, получаемых с камер, и данных, получаемых из модели.

В процессе исследования решены следующие задачи: создана программа автоматической обработки и моделирования данных о движении транспорта с интеллектуальной камеры; проведена оценка адекватности моделирования транспортных потоков на заданном перекрестке в условиях наличия полной информации о движении транспорта; проведена оценка адекватности моделирования транспортных потоков на заданном перекрестке при ограниченных объемах данных.

1. Исходные данные для моделирования транспортных потоков

Исходными данными для моделирования являются данные, получаемые с камер в виде файла в формате CSV⁴, в котором каждый проезд транспортного средства (далее ТС) через перекресток представляет собой ячейку таблицы. В нем хранится информация о типе проезжающего ТС, позиции старта и финиша, а также времени пересечения перекрестка. Для примера показана часть файла, представленная в виде табл. 1.

На основе таких данных может быть построена модель, воспроизводящая реальный трафик на дороге, которая может быть использована для построения прогноза изменения планов работы светофорных объектов.

⁴ CSV — Comma-separated values.

Таблица 1. Формат предоставленных данных

Тип ТС	Позиция старта	Позиция финиша	Дата и время
car	down	up	2023-06-30T09:00
car	up	down	2023-06-30T09:00:01
bus	left	down	2023-06-30T09:00:05
truck	left	up	2023-06-30T09:00:06

Входными данными такой модели, построенной в программе SUMO, являются позиция старта и финиша, а также дата и время записи информации о проезде.

Для сбора данных о трафике применяется интеллектуальная камера, использующая нейросетевые алгоритмы для обнаружения и классификации ТС, фиксации времени и направления проезда перекрестка. Предлагаемая программа в автоматическом режиме обрабатывает полученные с камеры данные, осуществляет подготовку к запуску модели в программе SUMO (где происходит непосредственно моделирование) и проводит анализ результатов моделирования. Схема взаимодействия модулей программы приведена на рис. 1. Программа позволяет автоматически без участия человека запускать процесс моделирования при поступлении новых данных с камеры.

Согласно схеме, исходные данные о реальном трафике, полученные с интеллектуальной камеры, считываются функцией *CSVReader* из *CSV* файла. Далее удаляются повторяющиеся данные о проезде транспортных средств, и функция *CountData* подсчитывает количество ТС за определенное время.

При помощи функции *ChangeTimeInPeriod* случайным образом определяется время появления каждой машины в симуляции. Время рассчитывается с использованием формулы:

$$A = \frac{T}{N} S + \Delta t, \quad (1)$$

где A — итоговое время появления ТС в моделировании в секундах;

T — период усреднения данных;

N — количество ТС, проезжающих перекресток за период T ;

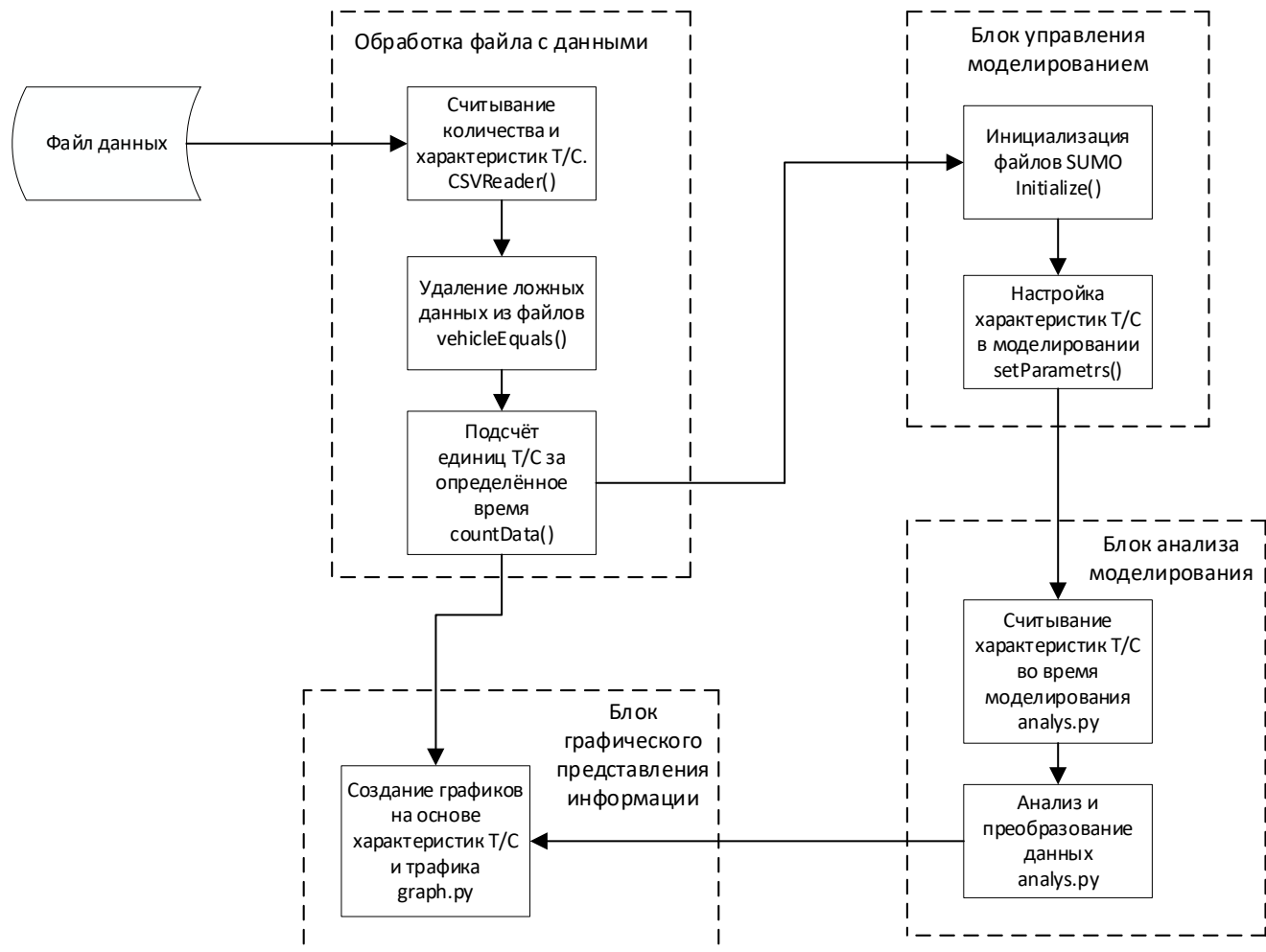


Рис. 1. Схема взаимодействия модулей

S — случайное число, подчиняющееся экспоненциальному [10] закону с параметром интенсивности $\lambda = 2,859$ 1/с;
 Δt — шаг моделирования;
 n — номер периода.

Параметр λ оценивается по текущим экспериментальным данным о транспортном потоке, получаемым с камеры, и может интерпретироваться как обратное среднее времени между появлением ТС [11]. Следовательно, оно может быть рассчитано как:

$$\lambda = \frac{1}{\delta}, \quad (2)$$

где δ — среднее время появления ТС в модели в секундах.

Далее определяется итоговое время появления ТС в симуляции путем суммирования

данных о рассчитанном случайном временном интервале и шаге моделирования.

Обработанные данные поступают в блок управления моделированием, где происходит настройка характеристик транспортных средств перед моделированием при помощи функции *setParams*, создание конфигурационных файлов для *SUMO* и вызов соответствующих команд для проведения моделирования при помощи функции *initialize*.

В процессе моделирования в программе осуществляются считывание и анализ таких основных характеристик как: скорость ТС, плотность и интенсивность движения.

У каждого транспортного средства в модели считывается текущая скорость, которая затем преобразуется в среднюю скорость движения транспортного потока. Она получается путем расчета среднего арифметического всех измеренных скоростей транспортных средств,



Рис. 2. Направления движения на перекрестке улиц Суворова — Володарского

находящихся в моделировании. Полученная таким образом средняя скорость потока является важным показателем эффективности транспортной системы. Она позволяет оценить общую скорость передвижения транспортных средств в заданной области и выявить возможные пробки или проблемные места на дорогах.

Важными аспектами симуляции в программе *SUMO* являются измерение плотности транспортного потока, как количества ТС, находящихся на дороге в единицу времени и определение интенсивности движения.

Обработанные данные поступают в блок графического представления информации, где происходит построение графиков зависимостей различных величин.

Таким образом, данная программа специализирована для работы с интеллектуальными камерами, что позволяет эффективно использовать данные с камер для имитационного моделирования транспортного потока на локальном участке дорожно-транспортной сети. Основное преимущество заключается в автоматизации процесса моделирования. Программа сама проводит все шаги: подготовку данных, запуск моделирования и обработку полученных статистических данных.

2. Моделирование одного перекрестка на основании реальных данных о трафике

Разработанную программу мы протестировали на примере перекрестка улиц

Суворова — Володарского города Пензы, что позволило установить соответствие ожидаемых и реальных данных.

Задание параметров транспортного потока в имитационной модели осуществляется на основе данных о его интенсивности. Оценка интенсивности производится путем подсчета количества ТС, проходящих через данное сечение дороги в единицу времени. Сечениями, на которых производится оценка, являются границы перекрестка. Для выбора временного интервала было проведено исследование.

В основе исследования лежит оценка схожести результатов моделирования реальным данным при различных интервалах оценки интенсивности дорожного движения для перекрестка, на котором установлена интеллектуальная камера [11], способная вести подсчет количества транспортных средств по направлениям движения, а также классифицировать 7 типов транспортных средств (рис. 2).

Схема перекрестка, представленного в модели, показана на рис. 3.

Выборка, состоящая из 7790 проездов машин, была использована для построения и оценки модели.

Результаты моделирования для различных интервалов оценивания приведены на графиках (рис. 4–6). На графике представлены две линии, отражающие общее количество ТС в модели. Сплошная линия показывает количество ТС, полученных по реальным данным, пунктирная — по результатам моделирования.

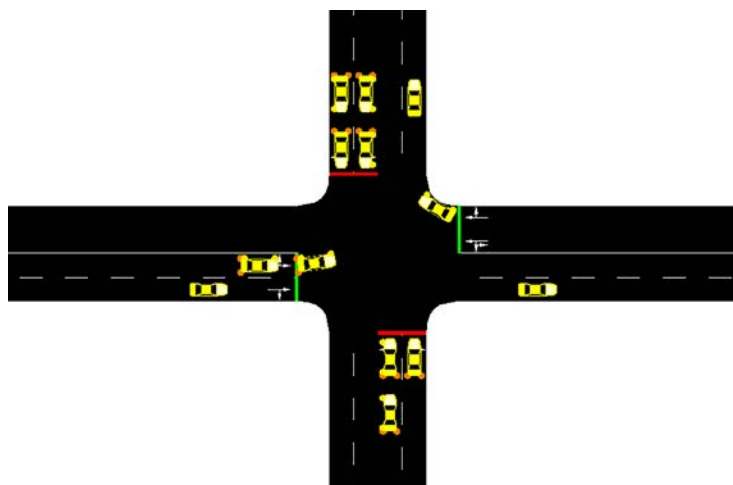


Рис. 3. Схема перекрестка

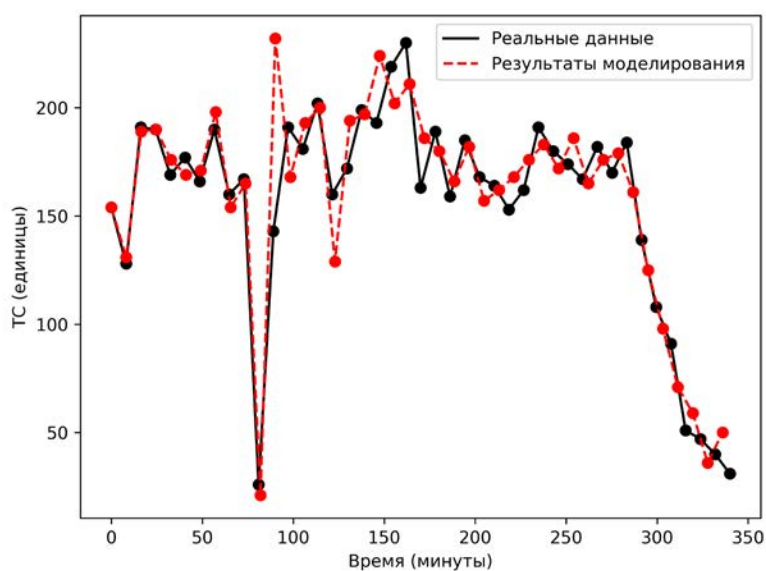


Рис. 4. Результаты моделирования для интервала оценивания 300 с

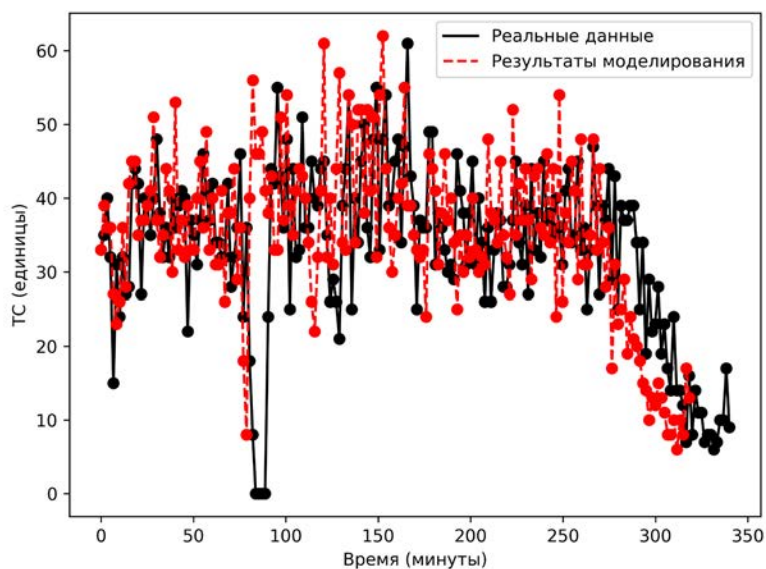


Рис. 5. Результаты моделирования для интервала оценивания 100 с

Можно сделать вывод, что при использовании слишком короткого интервала оценивания (100 с) график количества ТС, полученный в результате моделирования трафика, зашумляется из-за чувствительности к краткосрочным колебаниям в трафике. С другой стороны, при использовании слишком длинного интервала оценивания (1500 с) график становится слишком общим и не дает детального представления о потоке, что является важным для управления транспортным потоком на локальном участке дорожно-транспортной сети.

Для того чтобы выбрать наиболее подходящий интервал оценивания, предлагается использовать критерий корреляции Пирсона, рассчитывающийся по формуле [12]:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (3)$$

где x_i — значение интенсивности для конкретного интервала, вычисленное для реального трафика;

\bar{x} — среднее значение интенсивности, вычисленное для реального трафика;

y_i — значение интенсивности для конкретного интервала, вычисленное для смоделированного трафика;

\bar{y} — среднее значение интенсивности, вычисленное для смоделированного трафика.

Результаты расчета коэффициента корреляции показаны на рис. 7.

По графику видно, что наибольшее значение коэффициента корреляции на уровне 0,996 соответствует длительности 480 с. Это значение выбирается для вычисления интенсивности в последующих симуляциях. При выбранном периоде получили следующий график количества машин, представленный на рис. 8.

Интервал усреднения не оказывает влияния на плотность трафика. Перенос рекомендованной продолжительности данного периода на другие конфигурации перекрестков подразумевает анализ схожести в динамических особенностях трафика. Подразумевается анализ средней интенсивности движения в качестве ключевого фактора при переносе

рекомендованной продолжительности периода на другие конфигурации перекрестков.

Для полученных данных была рассчитана среднеквадратичная ошибка между реальными и смоделированными данными. Она составила 4,79 ТС.

Также для проведенного эксперимента рассчитаем доверительный интервал и стандартную ошибку. Расчет будет производиться, используя преобразование Фишера. Формула преобразования Фишера выглядит следующим образом [13]:

$$Z = \frac{1}{2} \ln \frac{1+r_{xy}}{1-r_{xy}}, \quad (4)$$

где Z — преобразованный коэффициент корреляции;

r_{xy} — исходный коэффициент корреляции. Таким образом, получаем $Z = 3,1063$.

Далее рассчитаем стандартную ошибку преобразованного коэффициента:

$$SE_Z = \frac{1}{\sqrt{n-3}}, \quad (5)$$

где n — количество наблюдений.

Получили стандартную ошибку $SE_Z = 0,1259$.

Для уровня доверительной вероятности $\gamma = 0,95$, принимаемой в большинстве инженерных расчетов [14], из соотношения $2\Phi(z_\gamma) = \gamma$ нашли коэффициент доверия z_γ , который равен 1,96.

Вычислили нижнюю и верхнюю границы интервала по формуле:

$$z_i = z \pm z_\gamma SE_z. \quad (6)$$

Получили верхнюю границу $z_1 = 3,345$ и нижнюю границу $z_2 = 2,855$.

Выполним обратные преобразования Фишера для возвращения размерности по формуле [13]:

$$r_i = \tanh(z_i). \quad (7)$$

Получили верхнюю границу $r_1 = 0,997$ и нижнюю границу $r_2 = 0,993$.

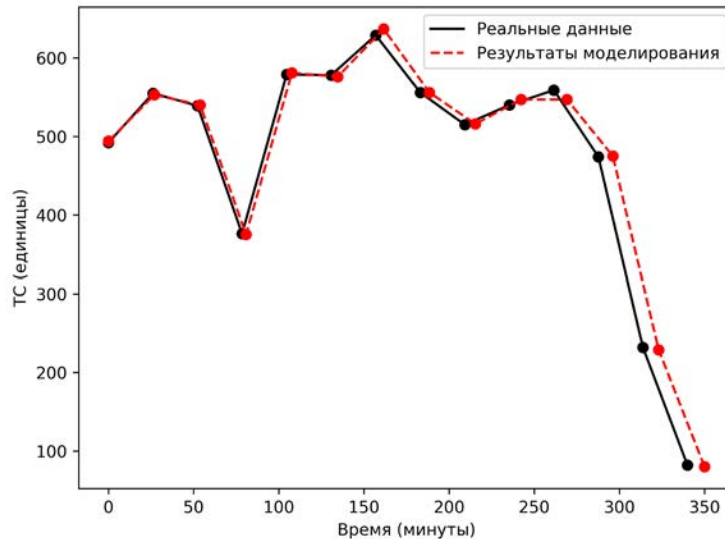


Рис. 6. Результаты моделирования для интервала оценивания 1500 с

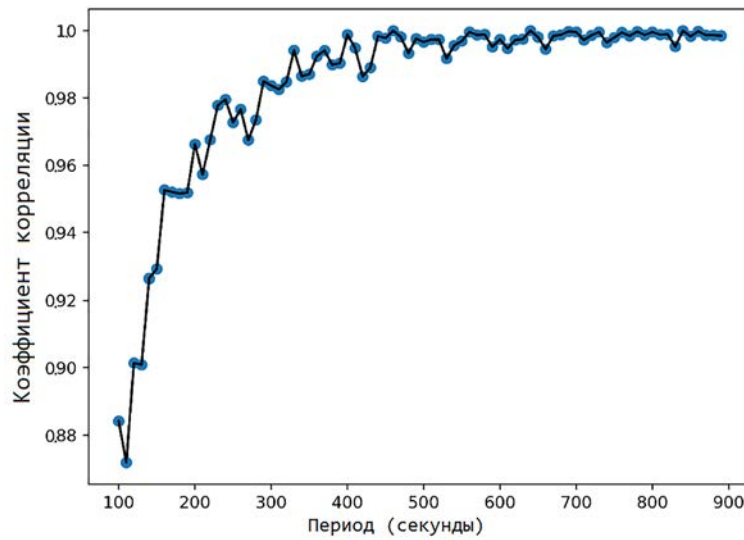


Рис. 7. График зависимости корреляции от длительности интервала оценивания

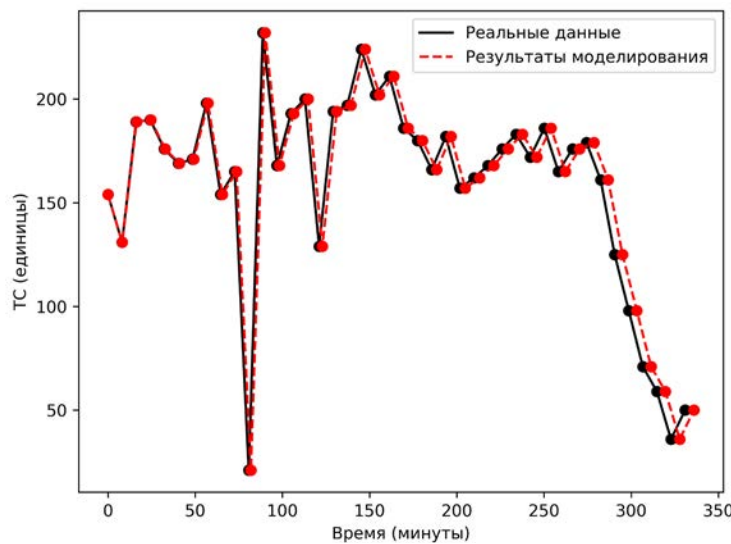


Рис. 8. Результаты моделирования для интервала оценивания 480 с

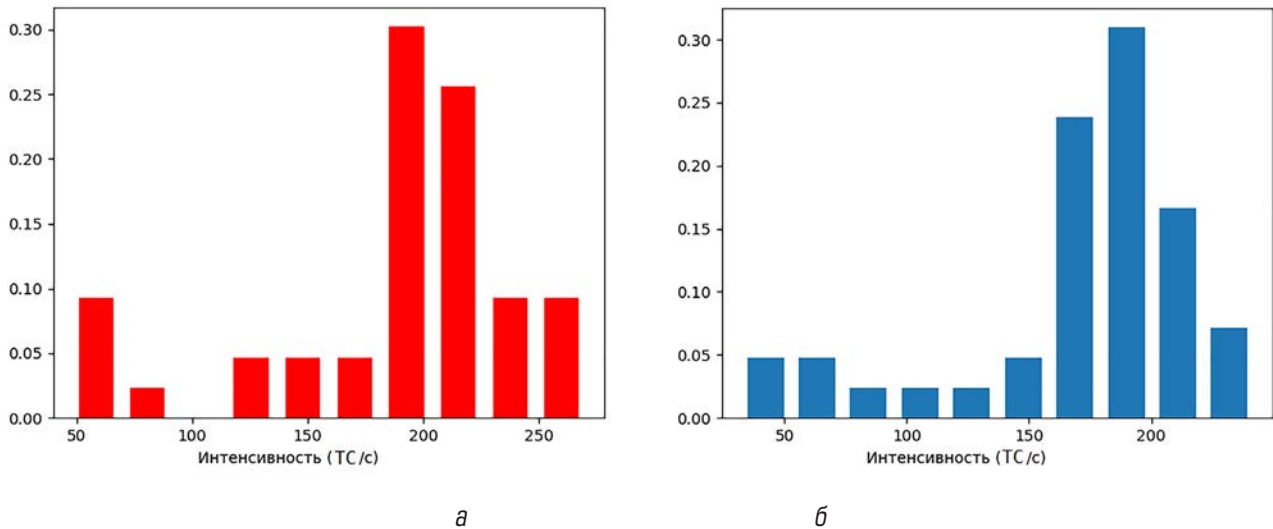


Рис. 9. Гистограммы интенсивности движения:

а — распределение интенсивности движения транспорта для интервала оценивания 480 с по реальным данным;
 б — распределение интенсивности движения транспорта для интервала оценивания 480 с по модельным данным

Таким образом, коэффициент корреляции $r = 0,996$ находится в пределах доверительного интервала ($r_1 < r < r_2$), это указывает на то, что данный коэффициент является статистически значимым на заданном уровне доверия.

На рис. 9, а показано распределение интенсивности движения транспорта для интервала оценивания 480 с по реальным данным, а на рис. 9, б — по модельным данным.

Для оценки различий между выборками интенсивностей, рассчитанных по реальным и смоделированным данным, использовался U -критерий Манна — Уитни, который подтвердил отсутствие статистически значимых различий на уровне значимости 0,1.

3. Моделирование на ограниченных наборах данных

Далее было проведено моделирование трафика для трех перекрестков на основании реальных данных о трафике через них. Это перекрестки улиц: Суворова — Плеханова, Суворова — Володарского и Суворова — Московская в г. Пензе.

В данном моделировании особое внимание уделяется возможности моделирования трафика, используя ограниченный набор данных. Смоделируем ситуацию, когда известен трафик только через два крайних перекрестка: Суворова — Плеханова и Суворова — Московская. Задача ставится как моделирование трафика

через третий перекресток по известным данным для двух других.

Данные для перекрестков представляют собой CSV-файл с записями о проездах транспортных средств в количестве 7487 строк для перекрестка Суворова — Московская; 13 846 строк для перекрестка Суворова — Плеханова и 13 845 строк для перекрестка Суворова — Володарского, где каждая строка представляет собой проезд машины по перекрестку. Низкий трафик на перекрестке улиц Володарского — Московская объясняется проведением ремонтно-восстановительных работ с последующим ограничением проезда на улице Московской.

При этом необходимо определить параметры распределения потока по направлениям на среднем перекрестке (Суворова — Володарского), основываясь на данных предварительного обследования. Мы используем метод расщепления потока, основанный на подсчете количества транспортных средств в каждом направлении в течение определенного периода времени. Затем мы вычисляем вероятность того, что транспортное средство повернет в определенном направлении, путем деления количества транспортных средств в данном направлении за указанный период на общее количество транспортных средств на перекрестке за этот период.

Схема перекрестка, реализованного в модели, представлена на рис. 10.

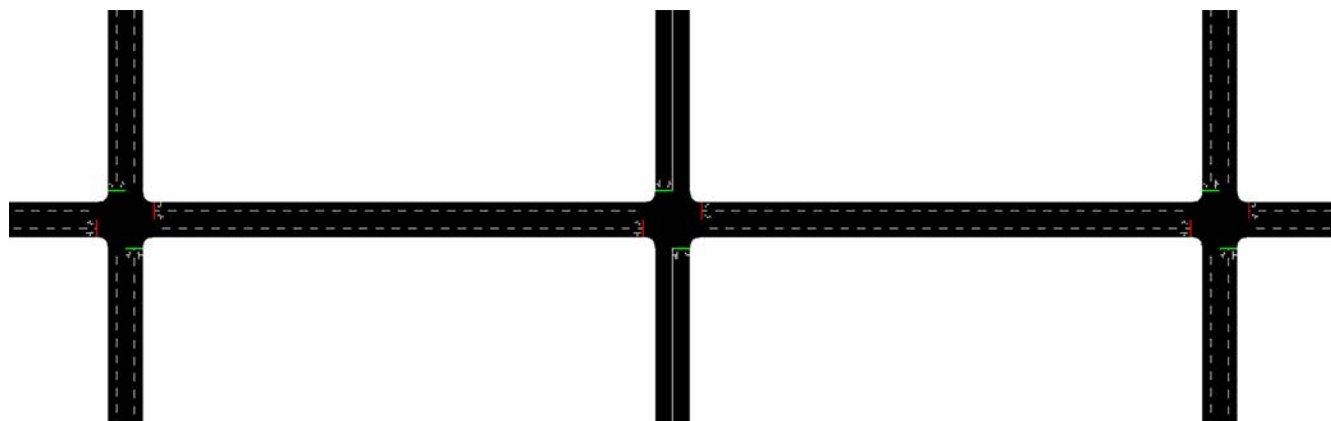


Рис. 10. Поле моделирования трех перекрестков

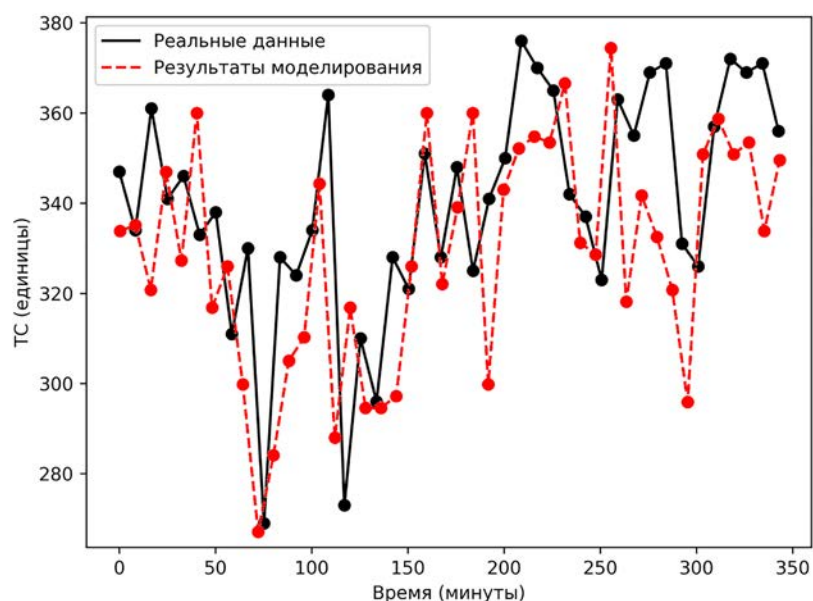


Рис. 11. Результаты моделирования на ограниченных наборах данных

При проведении моделирования с ограниченным объемом данных мы будем исключать информацию о движении на перекрестке улиц Володарского — Суворова из входного набора данных имитационной модели.

Для улицы Суворова данные о движении на перекрестке собираются с двух смежных перекрестков, в то время как в случае улицы Володарского мы предполагаем, что трафик схож с тем, что наблюдается на соседних улицах — Плеханова и Московской.

Практически все изучаемые перекрестки на улице Суворова проявили схожую динамику движения транспорта (в период с 9 до 15 ч количество проезжающих машин практически одинаково на всех перекрестках), что свидетельствует о стабильности движения в указанный

временной промежуток. На основании этой информации для моделирования на ограниченных наборах данных возьмем такой же период, как и при моделировании одного перекрестка.

Результаты моделирования на основе ограниченных наборов данных представлены на рис. 11. На графике сплошная линия показывает количество ТС, вычисленных по реальным данным, пунктирная линия — по результатам моделирования.

По результатам моделирования транспортных потоков через перекресток на ограниченных наборах данных среднеквадратичная ошибка для всего интервала моделирования составила 3,66 ТС, коэффициент корреляции между реальными и смоделированными данными r составил 0,7.

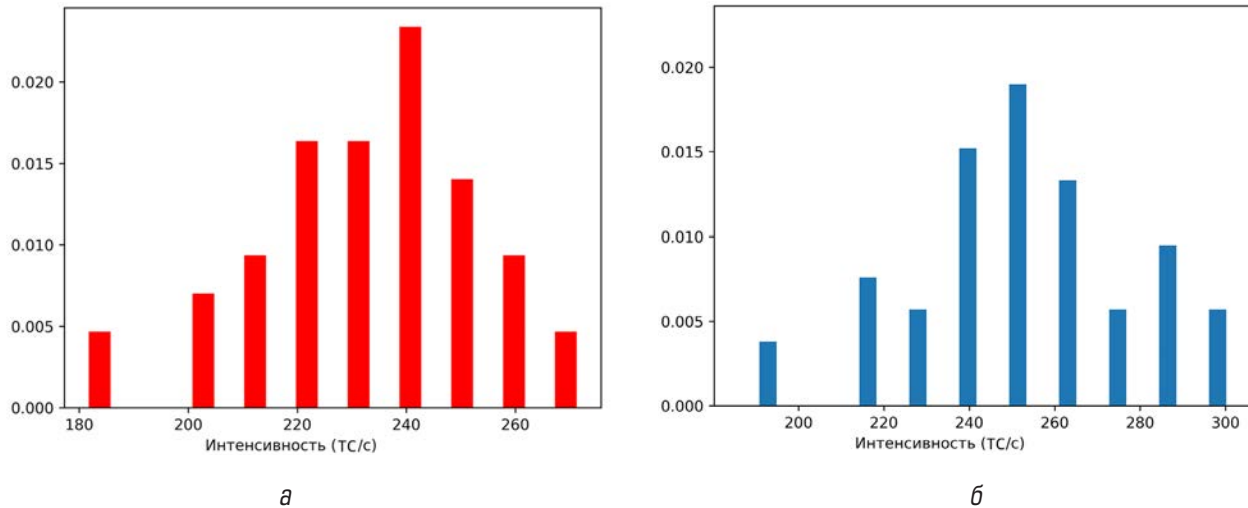


Рис. 12. Гистограммы интенсивности движения:
 а — распределение по реальным данным; б — по модельным данным

Также как и в случае с результатами моделирования одного перекрестка, применим преобразование Фишера к полученному коэффициенту корреляции, взяв в качестве доверительной вероятности $\gamma = 0,95$.

В результате вычислений получили доверительный интервал с верхней границей $r_1 = 0,824$, нижней границей $r_2 = 0,492$ и стандартную ошибку $SE_z = 0,162$.

Таким образом, коэффициент корреляции также находится в пределах доверительного интервала ($r_1 < r < r_2$), это указывает на то, что данный коэффициент является статистически значимым на заданном уровне доверия.

Также было проведено исследование распределения интенсивности движения транспорта для интервала оценивания 480 с. На рис. 12, а показано распределение по реальным данным, а на рис. 12, б — по модельным данным.

Для оценки различий между выборками интенсивностей, рассчитанных по реальным и смоделированным данным, также использовался U-критерий Манна — Уитни, который подтвердил отсутствие статистически значимых различий на уровне значимости 0,1.

4. Заключение

Основным научным результатом работы является модель, реализованная в виде программного комплекса, способного автоматически моделировать трафик без участия человека, обрабатывая данные, поступающие с

интеллектуальных камер, установленных на перекрестках.

В рамках проведенного исследования подтверждена возможность моделирования транспортных потоков в условиях неполноты исходных данных, а конкретно — возможность моделирования потоков на перекрестках, ориентируясь исключительно на информацию о движении транспорта в смежных областях.

Уместно подчеркнуть, что при моделировании конкретного перекрестка коэффициент корреляции достиг высоких 0,996, что указывает на значительную точность моделирования данного перекрестка.

Исследование подтверждает, что моделирование трафика на основе ограниченных данных в перспективе может стать инструментом в управлении транспортным движением. Среднеквадратичная ошибка моделирования в условиях неполных исходных данных составила 3,66 ТС, коэффициент корреляции между реальными и смоделированными данными составил 0,7, что соответствует средней степени совпадения результатов. Эти значения подтверждают умеренную точность модели, обеспечивая достаточное согласование между моделированием и фактическими данными.

В двух экспериментах коэффициенты корреляции находятся в диапазонах доверительных интервалов, что также говорит о достаточной сходимости результатов моделирования.

Результаты исследования имеют практическое значение с перспективой определения на модели оптимальных параметров управления транспортными потоками на локальных временных интервалах порядка нескольких десятков минут. Развитие модели планируется в направлении ее использования для создания адаптивных алгоритмов управления светофорными объектами на перекрестках. Этот шаг позволит повысить эффективность управления транспортной инфраструктурой к изменяющимся условиям дорожного движения.

Авторы благодарят ООО «КодИнсайд» за предоставленные данные и финансовую поддержку проведенного исследования. ▲

Библиографический список

1. Сальникова А. Д. Формирование системы транспортно-пересадочных узлов в городской агломерации / А. Д. Сальникова, А. В. Баните, Д. Г. Плотников и др. // Автоматика на транспорте. — 2023. — Т. 9. — № 1. — С. 87–98.
2. Макаров И. В. Совершенствование управления транспортными потоками города с использованием имитационного моделирования / И. В. Макаров, Р. Г. Хабибуллин, К. А. Шубенкова и др. // Материалы 4-й Всероссийской научно-практической конференции по имитационному моделированию ИММОД 2009. — СПб.: Дом ученых им. М. Горького. — 2009.
3. Shahrudin R. A. Controlling Traffic Congestion in Urbanised City: A Framework Using Agent-Based Modelling and Simulation Approach / R. A. Shahrudin, Md Yu. Misro // ISPRS Int. J. Geo-Inf. — 2023. — Iss. 12(6). — P. 226.
4. Alghamdi T. A. Comparative Study on Traffic Modeling Techniques for Predicting and Simulating Traffic Behavior / T. Alghamdi, S. Mostafi, G. Abdelkader et al. // Future Internet. — 2021. — Iss. 14. — P. 294.
5. Yedavalli P. Microsimulation analysis for network traffic assignment (MANTA) at metropolitan-scale for agile transportation planning / P. Yedavalli, K. Kumar, P. Waddell // Transportmetrica A: Transport Science. — 2021. — Pp. 1278–1299.
6. Кузнецов Т. А. Оптимизация движения транспортного потока на регулируемом перекрестке с использованием имитационного моделирования / Т. А. Кузнецов // Политехнический молодежный журнал. — 2022. — № 07(72).
7. Захаров Ю. И. Основные современные инструменты имитационного моделирования транспортных потоков / Ю. И. Захаров, Е. С. Карнаух // Вісник придніпровської державної академії будівництва та архітектури. — 2023. — С. 87–98.
8. Di Sh. Calibrating stochastic traffic simulation models for safety and operational measures based on vehicle conflict distributions obtained from aerial and traffic camera videos / Sh. Di, G. Jingqin, Y. Di et al. // Accident Analysis and Prevention. — 2023. — Iss. 179.
9. Ромашова А. А. Анализ технологий сбора данных о параметрах движения транспортных потоков / А. А. Ромашова // Форум молодых ученых. — 2021. — № 5(57). — С. 298–303.
10. Altshuler Ya. Augmented Betweenness Centrality for Mobility Prediction in Transportation Networks / Ya. Altshuler, R. Puzis, Yu. Elovici et al. // Finding Patterns of Human Behaviors in Network and Mobility Data. — 2011. — P. 671.
11. Ross S. M. Introduction to Probability and Statistics for Engineers and Scientists / S. M. Ross. — University of California, Berkeley. — 2004. — P. 179.
12. Rodgers J. L. Thirteen Ways to Look at the Correlation Coefficient / J. L. Rodgers, W. A. Nicewander // The American Statistician. — 1988. — Vol. 42. — Iss. 1. — Pp. 59–66.
13. Hotelling H. New Light on the Correlation Coefficient and its Transforms / H. Hotelling // Journal of the Royal Statistical Society, Series B (Methodological). — 1953. — Vol. 15(2). — Pp. 193–225.
14. Фишер П. А. Статистические методы для исследователей: монография / П. А. Фишер. — М.: Госстатиздат, 1958. — С. 42.

TRANSPORT AUTOMATION RESEARCH, 2023, Vol. 9, No. 4, pp. 355–367
DOI: 10.20295/2412-9186-2023-9-04-355-367

Modeling of Local Traffic Flows under Conditions of Incomplete Initial Data

Information about authors

Mitrokhin M. A., Doctor of Engineering, Professor. E-mail: mmax83@mail.ru

Alyayev A. O., Research Assistant. E-mail: alyayev-2003@mail.ru

Lobanov R. I., Student. E-mail: ramil_lobanov@mail.ru

Kiryayev V. E., Student. E-mail: vlad.kiryayev.2019@mail.ru

Penza State University, "Computer Engineering" Department, Penza

Abstract: The main scientific result of the work is a model implemented in the form of a software package capable of automatically modeling traffic without human intervention, processing data coming from intelligent cameras installed at intersections.

Within the framework of the study, the possibility of modeling traffic flows in conditions of incompleteness of the initial data has been confirmed, and specifically, the possibility of modeling flows at intersections, focusing solely on information about traffic in related areas.

The study confirms that traffic modeling based on limited data in the future can become a tool of traffic management. The mean squared error of modeling under conditions of incomplete initial data is 3,66, the correlation coefficient between real and simulated data is 0,7, which corresponds to the average degree of coincidence of the results.

It is appropriate to emphasize that, when modeling a particular intersection, the correlation coefficient has reached 0,996, which indicates a significant accuracy in modeling this intersection.

In two experiments, the correlation coefficients are in the ranges of confidence intervals, which also indicates sufficient convergence of the simulation results.

The results of the study are of practical importance with the prospect of determining the optimal parameters of traffic flow control on the model at local time intervals of approximately several tens of minutes. The development of the model is planned in the

direction of its use to create adaptive algorithms for controlling traffic light objects at intersections. This step will improve the efficiency of transport infrastructure management to changing traffic conditions.

Keywords: modeling of traffic flows; traffic; Simulation of Urban Mobility; data averaging interval.

References

1. Sal'nikova A. D., Banite A. V., Plotnikov D. G. Formirovanie sistemy transportno-perevadochnykh uzlov v gorodskoy aglomeratsii [Formation of a system of transport hubs in the urban agglomeration]. *Avtomatika na transporte* [Transport automation research]. 2023, vol. 9, Iss. 1, pp. 87–98. (In Russian)
2. Makarov I. V., Khabibullin R. G., Shubenkova K. A. et al. Sovershenstvovanie upravleniya transportnymi potokami goroda s ispol'zovaniem imitatsionnogo modelirovaniya [Improving the management of city traffic flows using simulation modeling]. *Materialy 4-y Vserossiyskoy nauchno-prakticheskoy konferentsii po imitatsionnomu modelirovaniyu IMMOD 2009* [Materials of the 4th All-Russian Scientific and Practical Conference on Simulation modeling IMMOD 2009]. St. Petersburg: Dom uchenykh im. M. Gor'kogo Publ., 2009. (In Russian)
3. Shaharuddin R. A., Misro Md Yu. Controlling Traffic Congestion in Urbanised City: A Framework Using Agent-Based Modelling and Simulation Approach. *ISPRS Int. J. Geo-Inf.*, 2023, Iss. 12(6), p. 226.
4. Alghamdi T. A., Mostafi S., Abdelkader G. et al. Comparative Study on Traffic Modeling Techniques for Predicting and Simulating Traffic Behavior. *Future Internet*, 2021, Iss. 14, p. 294.
5. Yedavalli P., Kumar K., Waddell P. Microsimulation analysis for network traffic assignment (MANTA) at metropolitan-scale for agile transportation planning. *Transportmetrica A: Transport Science*, 2021, pp. 1278–1299.
6. Kuznetsov T. A. Optimizatsiya dvizheniya transportnogo potoka na reguliruemom perekrestke s ispol'zovaniem imitatsionnogo modelirovaniya [Optimization of traffic flow at a controlled intersection using simulation modeling]. *Politekhnicheskiy molodezhnyy zhurnal* [Polytechnic youth magazine]. 2022, Iss. 07(72). (In Russian)
7. Zakharov Yu. I., Karnaukh E. S. Osnovnye sovremennye instrumenty imitatsionnogo modelirovaniya transportnykh potokov [Basic modern tools for simulation modeling of transport flows]. *Visnik pridniprovs'koi derzhavnoi akademii budivnitstva ta arkhitekturi* [Bulletin of the Dnieper State Academy of Civil Engineering and Architecture]. 2023, pp. 87–98. (In Russian)
8. Di Sh., Jingqin G., Di Y. et al. Calibrating stochastic traffic simulation models for safety and operational measures based on vehicle conflict distributions obtained from aerial and traffic camera videos. *Accident Analysis and Prevention*, 2023, Iss. 179.
9. Romashova A. A. Analiz tekhnologiy sbora dannykh o parametrah dvizheniya transportnykh potokov [Analysis of technologies for collecting data on the parameters of traffic flows]. *Forum molodykh uchenykh* [Forum of young scientists]. 2021, Iss. 5(57), pp. 298–303. (In Russian)
10. Altshuler Ya., Puzis R., Elovici Yu. et al. Augmented Betweenness Centrality for Mobility Prediction in Transportation Networks. *Finding Patterns of Human Behaviors in Network and Mobility Data*, 2011, p. 671.
11. Ross S. M. *Introduction to Probability and Statistics for Engineers and Scientists*. University of California, Berkeley, 2004, p. 179.
12. Rodgers J. L., Nicewander W. A. Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, 1988, vol. 42, Iss. 1, pp. 59–66.
13. Hotelling H. New Light on the Correlation Coefficient and its Transforms. *Journal of the Royal Statistical Society, Series B (Methodological)*, 1953, vol. 15(2), pp. 193–225.
14. Fisher R. A. *Statisticheskie metody dlya issledovateley: monografiya* [Statistical methods for researchers: monograph]. Moscow: Gosstatizdat Publ., 1958, p. 42. (In Russian)