

УДК 004.041

DOI: 10.25559/SITITO.15.201903.755-766

Об использовании данных мобильных абонентов в цифровой урбанистике

М. В. Булыгин*, Д. Е. Намиот

Московский государственный университет имени М. В. Ломоносова, г. Москва, Россия
119991, Россия, г. Москва, ГСП-1, Ленинские горы, д. 1

* messimm@yandex.ru

Аннотация

По исследованиям международного агентства "We are social" за 2018 год в мире насчитывается 5.135 миллиарда пользователей мобильных устройств, что составляет 68% от общей численности населения. В процессе функционирования каждого из мобильных устройств порождаются данные, которые могут быть использованы для решения прикладных задач. Статья состоит из двух частей. В первой части приведен краткий обзор методов обработки четырех видов данных, получаемых со смартфонов: данных сотового оператора, данных, полученных с акселерометров, GPS и Wi-Fi датчиков. Обзор каждой группы методов систематизирован в виде таблиц, которые помогут быстро найти метод для решения нужной задачи или для обработки имеющихся данных. Вторая часть посвящена выявлению основных преимуществ и недостатков использования каждого типа данных. Эта информация так же представлена в виде таблицы, которая позволит выбрать нужный тип данных для решения задачи. В работе приведен обзор статей, посвященных анализу основных недостатков и путей их преодоления, в частности, краудсенсингу. В заключении делается вывод о том, что дальнейшее развитие в области анализа данных, полученных со смартфонов, может состоять в решении существующих проблем, поиске новых приложений и совместном использовании разных типов данных.

Ключевые слова: данные мобильных абонентов, цифровая урбанистика, анализ данных.

Для цитирования: Булыгин М. В., Намиот Д. Е. Об использовании данных мобильных абонентов в цифровой урбанистике // Современные информационные технологии и ИТ-образование. 2019. Т. 15, № 3. С. 755-766. DOI: 10.25559/SITITO.15.201903.755-766

© Булыгин М. В., Намиот Д. Е., 2019



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



On the Use of Mobile Subscribers Data in Digital Urban Planning

M. V. Bulygin*, D. E. Namiot

Lomonosov Moscow State University, Moscow, Russia
1, Leninskie gory, Moscow 119991, Russia

*messimm@yandex.ru

Abstract

According to the research of the international agency "We are social" in 2018 there are 5.135 billion mobile users in the world, which is 68% of the total population. Each mobile device generates data that can be used for application purposes. The article consists of two parts. The first part provides a brief overview of how to process four types of data from smartphones: cellular operator data, accelerometer data, GPS and Wi-Fi sensors. The overview of each group of methods is systematized in the form of tables, which will help to quickly find a method to solve the desired task or to process the available data. The second part is devoted to identifying the main advantages and disadvantages of using each type of data. This information is also presented in the form of a table, which allows you to choose the right type of data for the task. This paper provides an overview of the articles on the analysis of the main shortcomings and ways to overcome them, in particular, crowdsourcing. The conclusion is that further development in the analysis of data obtained from smartphones may consist in solving existing problems, finding new applications and sharing different types of data.

Keywords: Mobile subscribers data, digital urbanism, data science.

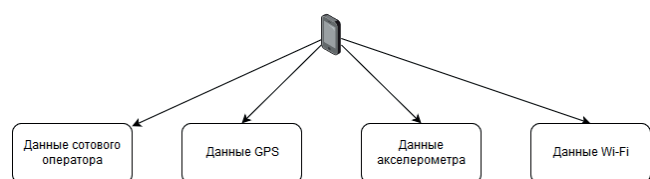
For citation: Bulygin M.V., Namiot D.E. On the Use of Mobile Subscribers Data in Digital Urban Planning. *Sovremennye informacionnye tehnologii i IT-obrazovanie* = Modern Information Technologies and IT-Education. 2019; 15(3):755-766. DOI: 10.25559/SITITO.15.201903.755-766



Введение

По исследованиям международного агентства “We are social” (2018) в мире насчитывается 5.135 миллиарда пользователей мобильных устройств, что составляет 68% от общей численности населения [1]. Прирост пользователей мобильных устройств по сравнению с состоянием 2017го года составил 218 миллионов пользователей, что составляет 4% от населения планеты.

В процессе использования мобильных устройств порождаются данные, которые могут быть использованы в прикладных задачах. В данной статье будут рассмотрены методы анализа четырех видов данных: данных сотовых операторов, данных с GPS-датчиков мобильных устройств, данные с акселерометров мобильных устройств, а также данные, полученные при помощи передатчиков Wi-Fi.



Р и с. 1. Данные, генерируемые смартфонами

Fig. 1. Smartphone generated data

Данные сотовых операторов

Данные о положении сотового телефона генерируются во время совершения звонков, отправки SMS-сообщений, при подключении к сети Интернет, а также при периодических обновлениях сотовой сети. Подробное описание стандартов и технологий, используемых при определении геопозиции можно найти в статье Ванга [2].

Использованию данных сотовых операторов для решения задач городского планирования посвящена статья «Urban Sensing Using Mobile Phone Network Data: A Survey of Research» Ф.Калабрезе, Л.Феррари и В. Блондела [3]. Исследователями предложено использовать эти данные для решения таких задач, как оценка распределения населения (оценка распределения проживания населения и оценка динамики), оценка типов активности в различных районах города, оценка паттернов мобильности, анализ локальных событий, анализ географии социальных сетей.

Авторы подробно описывают процесс генерации данных сотовой сети, как при действиях пользователей (отправка SMS-сообщений и звонки, передача Интернет-пакетов), так и при периодических обновлениях сети. Далее рассматриваются методы фильтрации данных: присвоение пользователю определенной геопозиции и обнаружение остановок. Исследователи рассматривают два типа данных: агрегированные по пространственным признакам и индивидуальные. Данные первого типа обладают меньшим объемом и могут быть доступны с временной задержкой чем данные второго типа, которые являются более подробными и позволяют работать с отдельными пользователями.

В другой работе Ф.Калабрезе «Real-Time Urban Monitoring Using Cell Phones: A Case Study in Rome» [4] показано, как можно использовать данные сотовых операторов на примере исследования, проведенного в Риме.

В начале статьи приведен краткий обзор и анализ ограничений существующих методов решения задач городского планирования, основанных на данных с GPS и сенсоров. Затем описывается платформа LocHNESSs, используемая для сбора данных о трафике. Авторы рассказывают о процессе преобработки данных: для оценки траекторий используется DFL-алгоритм, а для обнаружения остановок использованы пороговые значения скорости. Исследователями представлено шесть программных решений, отвечающих на следующие вопросы:

1. Где встречаются люди в Риме в течение дня?
2. Как люди перемещаются по определенным районам города во время мероприятий?
3. Где концентрируются иностранцы в Риме?
4. Как взаимно расположены транспорт и люди?
5. Куда движется транспорт?
6. Какие достопримечательности в Риме привлекают больше людей?

Одним из последних исследований в области использования данных сотовых операторов является работа С.Чена и др. «Data-Driven Prediction System of Dynamic People-Flow in Large Urban Network Using Cellular Probe Data». Авторы применили случайные леса к данным сотовых операторов. Для упрощения решаемой задачи исследователи отказались от решения задачи оценки точного количества людей в областях города (задачи регрессии), а решали задачу классификации. Модель предсказывает уровень загруженности каждой области. Первому уровню загруженности соответствуют значения от 0% до 20% от максимальной, второму – от 20% до 40%, и т.д. Шестому уровню загруженности соответствует загруженность выше 100% от максимальной наблюдаемой. Характерной особенностью подхода исследователей является наличие калибровки модели с учетом данных, поступающих с временной задержкой равной пяти минутам. Такая калибровка позволила значительно увеличить точность модели. Также авторы помимо данных сотовых операторов использовали данные о точках интереса, времени работы различных учреждений и данные о транспортной сети. Использование этих дополнительных данных также позитивно сказалось на качестве модели [5].

Исследователями во главе с Ф. Яном был предложен новый метод расчета времени в пути с использованием мобильных данных в статье «Arterial link travel time estimation considering traffic signal delays using cellular handoff data». Средние скорости движения рассчитываются путем деления предварительно откалиброванных расстояний между последовательными точками беспроводной передачи данных на промежутки времени. Предложенная модель сравнивается с другими существующими моделями, а также проводится эксперимент в китайском городе Чэнду, в котором модель демонстрирует ошибку времени прохождения артериального канала в основном в пределах 10% [6].

Новый метод поиска социальных событий был предложен Росарио Д. и др. в работе «Anomaly detection mechanisms to find social events using cellular traffic data». Исследователями представлены два подхода TA и SA на основе вейвлет-анализа совокупного сотового трафика, отличающихся тем, что они учитывают пространственную информацию о данных трафика. Учеными проведен вычислительный эксперимент на основе данных, полученных от итальянских операторов сотовой связи за декабрь 2013-го года по городу Милан. В результате



эксперимента были выявлены слабые стороны TA метода, выявлены особенности работы SA-метода, а также проведено сравнение с известным методом решения задачи, в результате которого оба представленных подхода показали более высокую производительность. В заключение авторами делается вывод, что стратегия, используемая в SA-методе, наилучшим образом подходит для обнаружений аномалий трафика, соответствующих локальным социальным событиям [7].

Также изучению коллективного поведения в больших масштабах и выявлению аномальных событий посвящена работа «Uncovering individual and collective human dynamics from mobile phone records» М.С. Гонзалез. Помимо изучения коллективного поведения в работе также представлены методы изучения моделей деятельности на индивидуальном уровне [8].

Оценке маршрутов на основе CDR посвящена работа «Cellpath Routing and Route Traffic Flow Estimation Based on Cellular Network Data». Исследователями предлагается несколько методов получения маршрутов на основе разреженных CDR-данных. Данные методы сравниваются на основе оценки при помощи них отдельных известных маршрутов. На основе эксперимента, проведенного на данных CDR для Дакара (Сенегал) авторы приходят к выводу, что выбор метода оценки маршрута может оказывать существенное влияние на результирующие маршруты [9].

Изучение социальных связей людей на основе CDR проведено в исследовании «The Relationship between Social Networks and Spatial Mobility: A Mobile-Phone-Based Study in Estonia». На основе эксперимента, проведенного в Эстонии, исследователи сделали выводы о том, что чем больше сеть партнеров, совершающих вызовы, а также чем больше зона, в которой живут вызывающие партнеры, тем больше перемещается человек. Также было установлено, что связь между звонками в сети партнеров и мобильностью зависит от пола [10]. Также социальные связи людей изучали Н.Игл и другие в работе «Inferring friendship network structure by using mobile phone data». Проанализировав данные, полученные со ста специально запрограммированных телефонов, ученые смогли выявить особенности, присущие дружеским отношениям, а также спрогнозировать результаты на индивидуальном уровне, такие, как удовлетворенность местом работы [11].

Работа Д.Ванга и др. «Human Mobility, Social Ties, and Link Prediction» посвящена изучению корреляции сходства траекторий движений с близостью в социальном графе. Используя данные о передвижении шести миллионов пользователей мобильной сети, ученые не только проследили данную корреляцию, но и исследовали возможности использования полученных данных для ответа на вопрос, какие новые связи будут развиваться в социальном графе [12].

Аналізу деятельности людей на основе мобильных данных посвящена работа К. Фаррахи и Д. Переса «Discovering routines from large-scale human locations using probabilistic topic models». Используя тематические модели, ученые смогли идентифицировать такие состояния, как «переход на рабочее место», «возвращение домой», «безостановочная работа», «рано уходить на работу» или «выключать мобильный по вечерам» на уровне пользователей и групп [13].

Мобильные данные также могут быть использованы и для картирования населения. Используя данные о вызовах, получаемые от сотовых операторов можно строить карты распределения населения почти в реальном времени. Один из методов картирования представлен в работе «Dynamic population mapping using mobile phone data» П.Девилья и др. Авторы показали возможности пространственной и временной оценки плотности населения на примере реальных данных для Португалии и Франции [14].

Новое решение проблемы размещения зон зеленых насаждений в городах было предложено командой исследователей во главе с Х.Бу в работе «Evaluation and Planning of Urban Green Space Distribution Based on Mobile Phone Data and Two-Step Floating Catchment Area Method». Авторы используют данные о местонахождении сотовых телефонов для учета фактических потребностей жителей в зеленых зонах, а также для анализа пропускной способности этих зон. Применяя 2SF-CA-алгоритм, исследователи выявили в пределах области исследования районы с низкой доступностью зеленых зон, которые можно рассматривать как ключевые для планирования зеленых зон в будущем. Авторами работы сделан вывод, что данные сотовых телефонов могут быть использованы как довольно точное отображение пространственного распределения жителей. Также отмечается, что методология, разработанная в рамках исследования, может быть применена для исследования распределения и использования других типов социальных объектов и сооружений [15].

В работе «Passenger Demand Prediction with Cellular Footprints» Дж.Чу и др. данные сотовой сети были использованы для предсказания потока пассажиров. Анализ набора данных сотовой связи, охватывающего 1,5 миллиона жителей крупного китайского города, показал сильную корреляцию между массовым потоком людей и пассажирским спросом. Исследователями была разработана система FlowFlexDP для предсказания пассажирского спроса, которая в отличие от существовавших решений учитывала не только историю заказов, но и информацию о потоке людей, извлеченную из сотовых данных. Для прогнозирования пассажирского спроса в районах произвольных форм и размеров система использует графическую сверточную нейронную сеть. Система была испытана на наборе данных DiDi Chuxing и показала лучший результат, чем современные системы прогнозирования [16].

Необычное применение мобильных данных было найдено группой исследователей во главе с М. М. Nyhan в работе «Quantifying population exposure to air pollution using individual mobility patterns inferred from mobile phone data». Авторы анализировали подверженность населения загрязнению воздуха частицами PM2.5. Исследователи использовали сотовые данные для обнаружения мест работы и проживания 407 435 пользователей мобильных телефонов, а затем с учетом этой информации сделали вывод, что игнорирование ежедневной мобильности потенциально может привести к неправильной классификации оценок воздействия на здоровье [17].

Обобщим информацию о рассмотренных статьях в виде таблицы:



Таблица 1. Задачи и методы (данные операторов)
Table 1. Tasks and methods (operator data)

Решаемая задача	Используемые методы	Исследование
Шесть различных задач городского планирования	Разработанный DFL-алгоритм, методы сравнения с пороговыми значениями, методы визуализации и картирования	«Real-Time Urban Monitoring Using Cell Phones: A Case Study in Rome» Ф.Калабрезе и др.
Предсказание уровня загруженности по областям города	Классификация на основе случайных лесов	«Data-Driven Prediction System of Dynamic People-Flow in Large Urban Network Using Cellular Probe Data» С.Чен и др.
Расчет времени в пути для абонента	Калибровка последовательных точек беспроводной передачи данных с последующим делением расстояния между точками на время между передачами для определения скорости	«Arterial link travel time estimation considering traffic signal delays using cellular handoff data» Ф.Ян и др.
Поиск социальных событий при помощи анализа мобильных данных	Вейвлет-анализ, TA и SA-алгоритмы	«Anomaly detection mechanisms to find social events using cellular traffic data» Д.Посарио и др.
Анализ индивидуальных и коллективных паттернов поведения	Статистические методы, анализ математических ожиданий и стандартных отклонений.	«Uncovering individual and collective human dynamics from mobile phone records» М.С. Гонсалес и др.
Сравнение методов вычисления маршрутов на основе CDR	Различные методы вычисления точек маршрута на основе CDR	«Cellpath Routing and Route Traffic Flow Estimation Based on Cellular Network Data» Н. Брейер
Исследование связи между социальными сетями и пространственной мобильностью	Статистические методы анализа	«The Relationship between Social Networks and Spatial Mobility: A Mobile-Phone-Based Study in Estonia». А.Пуупа
Выявление признаков, характерных для отношения дружбы	Непараметрическая процедура множественной регрессии квадратичного присвоения (MRQAP), факторный анализ	«Inferring friendship network structure by using mobile phone data» Н.Игл
Предсказание отношений в социальном графе	Корреляционный анализ	«Human Mobility, Social Ties, and Link Prediction» Д.Ванг
Выявление и идентификация различных видов человеческой активности	Тематическое моделирование	«Discovering routines from large-scale human locations using probabilistic topic models» К.Фаррахи и Д. Перес
Картирование населения	Диаграммы Вороного, регрессионные модели	«Dynamic population mapping using mobile phone data» П.Девиль и др
Поиск мест для размещения зеленых микрорайонов	2SFCA-алгоритм	«Evaluation and Planning of Urban Green Space Distribution Based on Mobile Phone Data and Two-Step Floating Catchment Area Method» Х.Бу
Предсказание транспортного потока пассажиров	Сверточные нейронные сети	«Passenger Demand Prediction with Cellular Footprints» Дж.Чу и др.
Анализ подверженности населения загрязнению воздуха частицами PM _{2.5}	Регрессионная модель, статистические методы и агрегация данных по районам	«Quantifying population exposure to air pollution using individual mobility patterns inferred from mobile phone data» М.М. Нихан

Отдельно необходимо отметить особенности обработки данных мобильных абонентов в Московском регионе (Россия). Глубокое проникновение мобильных технологий (например, в московском транспорте полностью работает как мобильная связь, так и мобильный Интернет) позволяет мобильным операторам собирать информацию о матрицах корреспонденций (в англоязычной литературе – OD matrix или origin-destination matrix для абонентов. Это информация о том, сколько человек в произвольный момент времени направилось из одного выделенного района в другой (например, с одной станции на другую). Соответственно, для транспортных задач, например, матрица корреспонденции (для московского метро, в частности) известна. В этом состоит отличие от классических подходов, где часто именно построение матрицы корреспонденции является основной задачей. Также, именно для московского региона характерно наличие агрегированных данных по перемещениям мобильных абонентов, которые собирает профильный департамент Правительства Москвы. Анализ возникающих

при этом задач, а также обзоры литературы есть в работах российских авторов [41-44]. Задачи, которые при этом решались – это кластеризация и анализ подобия временных рядов.

Данные GPS

Современный смартфон также содержит датчики для определения геопозиции при помощи GPS. Главным преимуществом данных этого типа является их достаточно высокая пространственная точность. Главным недостатком использования данных с датчиков GPS является их низкая массовость. Несмотря на то, что данные GPS могут быть получены не только со смартфонов, но и со специальных устройств, таких как навигаторы, их общее число ниже общего числа мобильных устройств[1][18]. Также важно отметить, что функции установления геопозиции используются непостоянно и не всеми пользователями (в отличие от данных сотовых операторов, которые генерируются всеми пользователями постоянно).



Рассмотрим несколько примеров использования данных GPS для решения задач городского планирования. В статье Х.Чанга «Detecting Taxi Travel Patterns using GPS Trajectory Data: A Case Study of Beijing» описан опыт использования данных GPS для анализа работы службы такси. Исследователями были посчитаны среднее время обслуживания, средняя скорость автомобилей в различные временные периоды. Авторами были выявлены четыре класса водителей такси по типу их предпочтений (работающие в центре города, работающие в пригородах, предпочитающие поездки на дальние расстояния и без предпочтений). Данная информация, по мнению исследователей, может оказаться полезной для повышения качества работы службы такси [19].

Также данные GPS были использованы С.Ху и др. в работе «Taxi Driver's Operation Behavior and Passengers' Demand Analysis Based on GPS Data». На основе данных, записанных в течение 204 рабочих часов таксистов в городе Шэньчжэнь и включающих в себя данные о времени, долготе, широте, скорости и информацию о том, находятся ли пассажиры в транспортном средстве такси. Были проанализированы данные о восьми транспортных зонах, а затем были изучены требования клиентов по местам отправления/назначения в разное время в будни и выходные дни. По мнению исследователей, эта работа поможет таксистам легче находить пассажиров, а пассажирам лучше планировать вызов такси [20].

Данные GPS использовались группой китайских исследователей для определения скорости движения грузовых автомобилей. Авторами работы «Truck Traffic Speed Prediction Under Non-Recurrent Congestion: Based on Optimized Deep Learning Algorithms and GPS Data» была применена GRU-модель, являющаяся улучшением рекуррентных нейронных сетей. Обучение нейронной сети проводилось с оптимизаторами Adam, Adadelta, SGD, RMSPROP. Вычислительные эксперименты проводились для четырех различных ситуаций: рабочий день, выходной день, день с ДТП, а также в дождливый день. Во всех ситуациях обученные модели показали результаты с ошибкой менее 6% [21].

В работе П. Гальперна и других «Assessing urban connectivity using volunteered mobile phone GPS locations» был проведен анализ выбора типа транспортного средства. 127 студентов в течение 6 лет предоставляли данные о местоположении со своих сотовых телефонов. К полученным данным был применен метод байесовской бета-регрессии для моделирования доли времени, проведенного в неактивном состоянии, при ходьбе и путешествии на скорости транспортного средства, относительно связности улиц и других экологических характеристик, измеренных в радиусе дома. В результате проведенного анализа было показано, что доля времени, потраченного на ходьбу, была выше в районах с большим количеством связанных улиц, а также в студенческих домах, которые были ближе к кампусу. В районах же с низкой связностью улиц учащиеся проводили больше времени в неактивном состоянии или в транспортных средствах. Исследователями был сделан вывод, что данные генерируемые мобильными телефонами с поддержкой GPS, могут обеспечить доступную и гибкую базу данных для поддержки проектирования городских районов [22].

Изучением распределения людей в лесопарках и зеленых зонах занималась команда исследователей во главе с Ю.Жа. В работе «Can trail spatial attributes predict trail use level in urban forest park? An examination integrating GPS data and space syntax theory» авторы использовали данные, полученные с GPS-трекеров с интервалом в 10 секунд. В общей сложности, исследователи получили 134 действительных образца в лесопарке Гонгцин и 210 действительных образцов в лесном парке Паотайван для анализа. Применив многофакторный регрессионный анализ и взяв визуальное соединение с водой, соединение с различными объектами при помощи троп, затененность и некоторые другие факторы в качестве ковариантных переменных, авторы смогли объяснить 31,7% ($p < 0,001$) и 27,3% ($p < 0,001$) различий в уровне использования территории парков Гонгцин и Паотайван. Результаты данной работы могут быть использованы менеджерами парка и дизайнерами для обеспечения для посетителей комфортных социальных условий [23].

В работе П.Чена и др. «A GPS data-based analysis of built environment influences on bicyclist route preferences» проанализировано влияние различных факторов на маршруты велосипедистов. Анализ был проведен на основе данных, собранных приложением для смартфонов CycleTracks при помощи GPS в городе Сиэтл. Используя две смешанные логит-модели, исследователи смогли выявить некоторые закономерности, наблюдаемые при выборе маршрута велосипедистами. Например, ими было установлено, что велосипедисты чаще выбирают ровные маршруты с медленным движением, а также маршруты с большим количеством фонарей и зеленых насаждений. Информация, которая может быть получена с помощью подобного исследования, может быть использована властями города для улучшения инфраструктуры для велосипедистов, что будет способствовать популяризации здорового образа жизни среди населения [24].

Проблеме анализа движения на перекрестках посвящена работа М.Деня и др. «Generating urban road intersection models from low-frequency GPS trajectory data». Исследователи предприняли попытку создать структурные модели перекрестков с использованием данных низкочастотных траекторий. Пространственный охват пересечений дорог сначала определяется на основе анализа горячих точек и кластеризации точек на основе триангуляции. Далее, улучшенный алгоритм кластеризации иерархических траекторий используется для адаптивного извлечения правил движения на перекрестке. Авторами был проведен вычислительный эксперимент, показавший высокую эффективность обработки низкочастотных нестабильных данных, а также извлечения структурных особенностей перекрестков. Методика, описанная в данной статье, является многообещающим решением в области извлечения данных о дорогах [25].

Установлением пространственно-временного распределения туристов в городе с использованием данных GPS занималась корейские исследователи во главе с Х.Ж. Юном. В своей работе «Spatiotemporal distribution of urban walking tourists by season using GPS-based smartphone application» они проанализировали пространственно-временные данные, собранные от 233 участников при помощи специального приложения для смартфонов, использующего данные GPS. С использованием инструментов ArcGIS авторы установили зависимость распределения туристов-пешеходов от сезона, также было установлено, что наибольшие постоянные скопления туристов наблюдаются на торговых площадях. Исследователи считают, что материалы их работы могут помочь планировщикам пешеходных туров [26].

Также данные GPS могут быть использованы для идентификации типа передвижения. Данной теме посвящена статья К.Стюарт и др. «A Review of GPS Trajectories Classification Based on Transportation Mode». Авторами рассмотрены существующие методы определения классификации траекторий на основе режима транспортировки. Методы разбиты на две группы:



методы первой группы позволяют определять класс только для траекторий, которые записаны при одном типе движения, а методы второй группы способны распознавать эпизоды активности и точки перехода. Также авторами выделены существующие проблемы и ограничения применения данных методов. Материалы данной статьи могут быть использованы исследователями для выбора подходящего метода классификации траекторий для решения конкретной задачи [27].

Новый метод предсказания транспортных потоков был предложен Ж.Дуаном и др. в работе «Improved Deep Hybrid Networks for Urban Traffic Flow Prediction Using Trajectory Data». В ка-

честве исходных данных исследователями были использованы данные о GPS-траекториях водителей такси. Предложенный подход объединяет сверточную нейронную сеть с рекуррентной LSTM-сетью. Для сокращения времени обучения был использован жадный подход к обучению. Вычислительный эксперимент с реальными данными о траекториях такси в городе Сиань показал, что предложенная модель на основе глубокого обучения обеспечивает более высокую точность и меньшее время прогнозирования, чем другие существующие подходы [28].

Обобщим обзор в виде таблицы:

Таблица 2. Задачи и методы (GPS)
Table 2. Tasks and methods (GPS)

Решаемая задача	Используемые методы	Исследование
Расчет среднего времени обслуживания, средняя скорость автомобилей в различные временные периоды для службы такси	Анализ скрытых классов, методы кластеризации (метод k-средних)	«Detecting Taxi Travel Patterns using GPS Trajectory Data: A Case Study of Beijing» Х.Чанг
Исследование предпочтений пассажиров и водителей такси	Статистические методы анализа	«Taxi Driver's Operation Behavior and Passengers' Demand Analysis Based on GPS Data» Ж.Ху
Определение скорости движения грузовых автомобилей в потоке	GRU-модели с различными оптимизаторами (RMSPROP, SGD, Adam, Adadelta)	«Truck Traffic Speed Prediction Under Non-Recurrent Congestion: Based on Optimized Deep Learning Algorithms and GPS Data» Д.Жао
Изучение влияния планировки улиц на активность студентов	Байесовская бета-регрессия	«Assessing urban connectivity using volunteered mobile phone GPS locations» П.Гальперн и др.
Изучение распределения людей в лесопарках и выявление зон наибольшего интереса	Многофакторный регрессионный анализ	«Can trail spatial attributes predict trail use level in urban forest park? An examination integrating GPS data and space syntax theory» Ю.Жа
Изучение различных факторов, влияющих на маршруты велосипедистов	Смешанные логит-модели	«A GPS data-based analysis of built environment influences on bicyclist route preferences» П.Чен
Создание структурных моделей перекрестков	Иерархические методы кластеризации	«Generating urban road intersection models from low-frequency GPS trajectory data» М.День
Установление пространственно-временного распределения туристов-пешеходов	Статистические методы анализа, инструменты ArcGIS	«Spatiotemporal distribution of urban walking tourists by season using GPS-based smartphone application» Х.Ж. Юн
Классификация траекторий пользователей по типу передвижения	Рассмотрены различные методы классификации, способные определять класс только для траекторий, которые записаны при одном типе движения, так и методы, способные распознавать эпизоды активности и точки перехода	«A Review of GPS Trajectories Classification Based on Transportation Mode» К.Стюарт
Предсказание транспортных потоков	Гибридная нейронная сеть (LSTM+CNN)	«Improved Deep Hybrid Networks for Urban Traffic Flow Prediction Using Trajectory Data» Ж.Дуан

Данные, полученные с акселерометров

Современные смартфоны также оснащаются акселерометрами и гироскопами. Данные с этих датчиков также могут использоваться для решения задач городского планирования.

Показания акселерометров мобильных телефонов были использованы группой исследователей во главе с Т.Брезмесом в работе «Activity Recognition from Accelerometer Data on a Mobile Phone». Данные с акселерометра в данной работе были представлены трехкомпонентными векторами, в которых содержалась информация о силах, действующих на смартфон вдоль трех осей (oX, oY, oZ), Применяя метод k-ближайших соседей, авторы работы смогли добиться высокой точности распознавания таких действий, как ходьба, спуск и подъем по лестнице, падения и посадка [29].

Movement	1 train session	2 train sessions	Full train
Walking	70%	70%	90%
Climbing-down stairs	20%	20%	80%
Climbing-up stairs	60%	60%	80%
Sitting down	30%	30%	70%
Standing up	50%	50%	70%
falling	70%	80%	90%

Р и с. 2. Результаты применения метода k-ближайших соседей
Fig. 2. Results of the k-nearest neighbors method

Также хорошую точность распознавания активности удалось получить китайским ученым во главе с Х.Чжао. Исследователи применяли нейронные глубокие Bi-LTSM-сети для обработки данных с акселерометров и гироскопов смартфонов. Авторами работы распознавалось шесть видов двигательной активности: поездки на велосипеде, ходьба, бег, поездки на авто-



буса, остановки, поездки на метро. Для сбора данных были использованы смартфоны Xiaomi и Huawei P9. В статье представлены методы сбора данных, их обработки, извлечения признаков и их нормализации для последующего использования. Сравним эффективность применения различных архитектур нейронных сетей (в качестве показателей качества выступали precision, recall и F1-метрика), исследователями был сделан вывод, что наилучшей архитектурой нейронной сети для решения задачи является Bi-LSTM, показавшая значение метрики F1 более 0.92.

Авторы работы отмечают, что они и дальше продолжают совершенствовать результаты, расширяя набор данных для исследования и увеличивая точность распознавания [30]. Задача распознавания видов активности по данным акселерометра была решена множеством исследователей. А.Игнатовым был предложен подход на основе использования сверточных нейронных сетей. Главной особенностью работы является исследование зависимости точности предсказания от длины временного ряда, в результате которого был сделан вывод о возможности ограничения его до одной секунды, что позволяет вести классификацию деятельности в реальном времени. Решение, сконструированное в работе, было протестировано

	Predicted Class						
	BIC	BUS	RUN	STI	SUB	WAL	Recall
BIC	999	36	13	0	3	6	0.945
BUS	9	466	0	0	23	0	0.936
RUN	0	0	758	0	0	0	1
STI	0	0	0	548	184	0	0.749
SUB	2	0	0	6	682	0	0.991
WAL	23	3	0	0	0	480	0.949
Precision	0.969	0.923	0.983	0.989	0.765	0.988	
F1	0.957	0.929	0.991	0.852	0.864	0.968	0.928

Рис. 3. Показатели качества распознавания типов двигательной активности

Fig. 3. Indicators of the quality of recognition of types of locomotor activity

на наборах WISDM и UCI и показало эффективность на уровне современных систем [31].

Данные об этих и других исследованиях в области распознавания типов активности по данным акселерометров представим в виде таблицы:

Таблица 3. Задачи и методы (акселерометр)
Table 3. Tasks and methods (accelerometer)

Исходные данные	Распознаваемые типы активности	Используемые методы	Исследование
Данные с акселерометров мобильных телефонов	Ходьба, спуск и подъем по лестнице, сидение, вставание, падение.	Метод k-ближайших соседей	«Activity Recognition from Accelerometer Data on a Mobile Phone» Т.Брезмес
Данные с акселерометров и гироскопов смартфонов Xiaomi и Huawei P9	Поездки на велосипеде, ходьба, бег, поездки на автобусе, поездки на метро, остановки	глубокие Bi-LSTM-сети	Recognition of «Transportation State by Smartphone Sensors Using Deep Bi-LSTM Neural Network» Х.Чжао
Показания акселерометров, записанные с интервалом 50 мс.	Ходьба, бег, подъем и спуск по лестнице, сидение, вставание.	Сверточные нейронные сети	«Real-time human activity recognition from accelerometer data using Convolutional Neural Networks» А. Игнатов, «A robust convolutional neural network for on-line smartphone-based human activity recognition» А. Бандар и др.[32]
Показания акселерометров со смартфонов	Бег, ходьба, сидение, вставание	Случайные леса и градиентный бустинг (проведено сравнение методов)	«Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results» К. Ли [33]
Общедоступный набор данных компании HTC, включающий в себя данные с акселерометров, гироскопов и магнитометров смартфонов.	Распознавание транспортного средства: автомобиль, мотоцикл, автобус, велосипед, ходьба. Распознавание действия: повороты влево, вправо, движение по прямой, остановки	Случайные леса, наивный Байес, дерево решений, метод k-ближайших соседей, SVM.	«Vehicle Mode and Driving Activity Detection Based on Analyzing Sensor Data of Smartphones», Д.Н. Лю и др. [34]

Данные Wi-Fi-передатчиков

Современные смартфоны имеют Wi-Fi передатчики. Данные с них могут быть использованы как для определения гео-позиции и скорости, так и для определения типов мобильности на различных участках. Определение местоположения пользователя происходит при помощи Wi-fi positioning system. Работа данной системы основана на анализе ближайших точек

беспроводного доступа с известным местоположением, а также на их уровне сигнала.

Для определения местоположения пользователя по Wi-Fi может быть использована технология Google Services Geofences, которая определяет вхождение пользователя в некоторый радиус. Данная технология использует помимо Wi-Fi сигнала и информацию о сотовых вышках. В России похожую функциональность предоставляет сервис Яндекс.Локатор. Он



имеет открытый API и позволяет определять местоположение пользователя по MAC-адресам точек доступа, IP-адресу и ближайшим сотовым вышкам.

Другие работы посвящены анализу трафика при помощи статичных приемников/передатчиков сигнала. Команда исследователей под руководством М.Вона предложила решение задачи распознавания трафика на основе анализа данных Wi-Fi передатчиков. Работа их системы основана на том факте, что при прохождении объекта между приемником и передатчиком происходит изменение характеристик сигнала сети. Данные для исследования собирались при помощи двух ноутбуков, установленных на противоположных сторонах дороги. Для подсчета реальных данных, используемых в процессе обучения и тестирования, также были использованы видеорекамеры, при этом количество трафика и его тип подсчитывались вручную. Данные с передатчиков ноутбуков предварительно обрабатывались путем их перевода в изображения. Полученные изображения подавались на вход сверточной нейронной сети. Обученная нейронная сеть показала точность распознавания проезда транспортного средства с точностью 99,4%. Точность распознавания типа трафика в среднем составила 91,1% [35].

Группа ученых во главе с Я.Мостофи продемонстрировала подход, при помощи которого возможно подсчитать количество пешеходов. Исследователи установили два приемника/передатчика сигнала Wi-Fi на двух сторонах площадки размером 70 квадратных метров. После обработки данных и проведения вычислительного эксперимента авторы добились точности распознавания пешеходов 92% вне помещения и 88% внутри помещения [36].

Недостатки использования данных, полученных со смартфонов: мобильные данные

В работах [3][4][5] исследователями были выделены такие преимущества использования сотовых данных, как актуальность данных, их сравнительно невысокая стоимость (в отличие, например, от переписей населения, которые являются дорогостоящими и проводятся редко).

В работе [3] авторами были выделены и некоторые проблемы при работе с данными сотовых операторов. Например, информация о местоположении записывается с необходимой для исследований частотой только для достаточно активных пользователей, так как без звонков и отправки SMS-сообщений информация о пользователе записывается только во время периодических обновлений сети. Частично решить данную проблему возможно, если учитывать данные, генерируемые и при передаче пакетных данных, однако данной услугой пользуются лишь владельцы смартфонов, степень использования которых существенно ниже [1].

Другой проблемой, выделенной в этой же статье, является низкая пространственная точность данных. Существующие технологии и стандарты, описанные в [2], позволяют определить лишь приближенную геопозицию пользователя и не могут быть использованы для решения задач, требующих высокой пространственной точности (например, задачу определения загруженности конкретной улицы). Кроме того, является важным вопрос анонимности при анализе данных индивидуального типа. В работе Де Монтойе «The privacy

bounds of human mobility» было показано, что 95% пользователей возможно идентифицировать по четырем пространственно-временным точкам [37]. В индивидуальных CDR активных пользователей содержится более четырех записей, что может привести к возможности отслеживания пользователя без его согласия по данным его последующих CDR.

Проблеме конфиденциальности в современных сотовых сетях посвящена работа J.Rotha и др. «Quantifying Location Privacy in Urban Next-Generation Cellular Networks». Исследование иллюстрирует возможные пути атак на основе определения местоположения при помощи данных современных сотовых сетей, в которых управляющая информация передается без шифрования. Авторами определена количественная эффективность такой атаки, как в обычной сотовой среде, так и в городской среде. Исследователи показали возможность достижения точности атаки порядка 15 метров [38].

Недостатки использования данных, полученных со смартфонов: данные GPS

В рассмотренных выше исследованиях данные были получены от участников движения определенных классов, которые были обязаны предоставлять данные GPS со своих устройств при помощи приложений. Выделять же данные определенных классов участников движения в случае анализа мобильных данных затруднительно в виду их структуры, описанной в [2] и [3].

Чтобы решать прикладные задачи не только для отдельных классов, необходимо увеличить степень проникновения GPS-устройств. В качестве решения данной проблемы могут быть использованы специальные приложения, собирающие данные о геопозиции пользователя. Для мотивации пользователей данного приложения возможно использование финансовых поощрений или бонусов у компаний партнеров. Отметим, что пространственные данные, полученные путем обработки данных сотовых операторов по структуре похожи на данные GPS, поэтому возможно применение аналитических методов, описанных в этих статьях к данным GPS. Возможно, это позитивно скажется на результатах в виду более высокой точности, но эффект может и не проявиться по причине меньшей степени проникновения и использования GPS.

Недостатки использования данных, полученных со смартфонов: данные с акселерометров

В случае использования данных с акселерометров в больших масштабах можно получить данные о том, как движутся люди, а дополнив эту информацию данными с GPS-датчиков, возможно получение информации о том, какова загруженность дорог, пешеходных улиц, общественного транспорта и степень его использования и т.д. Основную сложность в данном подходе представляет получение данных. Так как они сохраняются на стороне пользователей, то это потребует от них установки отдельного приложения, которое в случае большой частоты сбора данных будет усиленно снижать заряд батареи устройства.



Недостатки использования данных, полученных со смартфонов: данные Wi-Fi

Данные о геопозиции, полученные с помощью Wi-Fi передатчиков, могут быть использованы аналогично GPS данным и данным от сотовых операторов. Использование Wi-Fi сетей превышает использование GPS, а пространственная точность такого определения геопозиции может оказаться выше, чем у методов с использованием данных сотовых операторов (так как помимо этих данных используются данные о точках беспроводного доступа). Главной проблемой при работе с такими данными является то, что данные о геопозиции сохраня-

ются на стороне пользователя и для их получения необходимо, чтобы пользователь установил приложение и разрешил использование его данных.

Подход, описанный в рассмотренных выше статьях, для использования на мобильных устройствах нуждается в развитии. В данных статьях описывалось два статичных приемника/передатчика сигнала, в то время как в случае с мобильными устройствами как минимум одно из двух взаимодействующих устройств движется. Проблема сохранения на стороне пользователя также существует при использовании данных этого типа.

Достоинства и недостатки описанных типов данных представим в виде таблицы:

Таблица 4. Достоинства и недостатки типов данных
Table 4. Advantages and disadvantages of data types

Тип данных	Преимущества	Недостатки
Данные сотовых операторов	Широкое распространение (собираются с каждого мобильного телефона), простота получения	Низкое пространственное и временное разрешение.
Данные GPS	Высокая пространственная точность, возможность более точного определения скорости	Трудность получения. Для сбора требуется специальное приложение
Данные акселерометров	Обеспечивают высокую точность решения задач классификации действий	Трудность получения. Для сбора требуется специальное приложение. Круг задач, решаемых с их использованием уже, так как акселерометр не оценивает положение в пространстве
Данные Wi-Fi	Широкое распространение, массовость использования.	Невысокая пространственная точность. Нужно стороннее ПО для сбора данных

Решением проблемы со сбором данных с датчиков GPS, акселерометров и Wi-Fi передатчиков может служить краудсенсинг. Краудсенсинг (англ. crowdsensing) позволяет сотовому оператору анонимно собирать информацию с датчиков пользователей и далее использовать ее для мониторинга, планирования и различных сервисов. В частности, О. Галининой был представлен подход обмена данных с датчиков пользователей на энергию. Это должно простимулировать интерес пользователей к тому, чтобы делиться данными, собранными датчиками их смартфонов [39].

Одной из главных проблем краудсенсинга являются высокие затраты энергии при сборе и отправке информации. Существует множество исследований по данной теме. Одним из них является работа N.D. Lane и др. «Piggyback CrowdSensing (PCS): energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities», посвященная энергоэффективному сбору данных с датчиков Wi-Fi и GPS. Главной особенностью предложенной системы является сбор данных во время активности пользователя, что позволяет снизить затраты электроэнергии за счет отсутствия необходимости выводить устройство из спящего режима [40]. Есть серия работ, посвященных back-end сервисам для задач краудсенсинга [45-47].

Заключение

Подходы, представленные в данной статье, используют данные, получаемые с мобильных устройств для решения задач городского планирования. Дальнейшее развитие в этой области может состоять в решении существующих проблем для данных описанных типов, в использовании комбинаций

различных типов данных с мобильных устройств, а также в поиске новых методов решения задач городского планирования, базирующихся на этих данных.

References

- [1] Global Digital Report 2018. Available at: <https://digitalreport.wearesocial.com/> (accessed 16.05.2019). (In Eng.)
- [2] Wang S., Min J., Yi B.K. Location-Based Services Mobiles: Technologies and Standards. Tutorial given at IEEE ICC 2008, Beijing, 2008. (In Eng.)
- [3] Calabrese F., Ferrari L., Blondel V.D. Urban Sensing Using Mobile Phone Network Data: A Survey of Research. *ACM Computing Surveys*. 2014; 47(2):25. (In Eng.) DOI: 10.1145/2655691
- [4] Chen X. et al. Data-Driven Prediction System of Dynamic People-Flow in Large Urban Network Using Cellular Probe Data. *Journal of Advanced Transportation*. 2019; 2019:9401630. (In Eng.) DOI: 10.1155/2019/9401630
- [5] Yang F., Yao Z., Jin P.J., Xiong Y. Arterial link travel time estimation considering traffic signal delays using cellular hand-off data. *IET Intelligent Transport Systems*. 2019; 13(3):461-468. (In Eng.) DOI: 10.1049/iet-its.2018.5186
- [6] Roth J., Tummala M., McEachen J., Scrofani J. Quantifying Location Privacy in Urban Next-Generation Cellular Networks. In: *Proceedings of the 51st Hawaii International Conference on System Sciences*. 2018, pp. 5831-5838. (In Eng.) DOI: 10.24251/hicss.2018.731
- [7] Candia J., Gonzalez M.C., Wang P., Schoenharl T., Madey G., Barabasi A-L. Uncovering individual and collective human dynamics from mobile phone records. *Journal of Physics A:*



- Mathematical and Theoretical*. 2007; 41(22):224015. (In Eng.) DOI: 10.1088/1751-8113/41/22/224015
- [8] Breyer N., Gundlegård D., Rydergren C. Cellpath Routing and Route Traffic Flow Estimation Based on Cellular Network Data. *Journal of Urban Technology*. 2018; 25(2):85-104. (In Eng.) DOI: 10.1080/10630732.2017.1386939
- [9] Puura A., Silm S., Ahas R. The Relationship between Social Networks and Spatial Mobility: A Mobile-Phone-Based Study in Estonia. *Journal of Urban Technology*. 2018; 25(2):7-25. (In Eng.) DOI: 10.1080/10630732.2017.1406253
- [10] Eagle N., Pentland A., Lazer D. Inferring Friendship Network Structure by Using Mobile Phone Data. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106(36):15274-15278. (In Eng.) DOI: 10.1073/pnas.0900282106
- [11] Wang D., Pedreschi D., Song C., Giannotti F., Barabasi A-L. Human mobility, social ties, and link prediction. In: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '11)*. Association for Computing Machinery, New York, NY, USA, 2011, pp. 1100-1108. (In Eng.) DOI: 10.1145/2020408.2020581
- [12] Farrahi K., Gatica-Perez D. Discovering routines from large-scale human locations using probabilistic topic models. *ACM Transactions on Intelligent Systems and Technology*. 2011; 2(1):3. (In Eng.) DOI: 10.1145/1889681.1889684
- [13] Deville P., Linard C., Martin S., Gilbert M., Stevens F.R., Gaughan A.E., Blondel V.D., Tatem A.J. Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*. 2014. 6 p. (In Eng.) DOI: 10.1073/pnas.1408439111
- [14] Wu H., Liu L., Yu Y., Peng Z. Evaluation and Planning of Urban Green Space Distribution Based on Mobile Phone Data and Two-Step Floating Catchment Area Method. *Sustainability*. 2018; 10(1):214. (In Eng.) DOI: 10.3390/su10010214
- [15] Chu J. et al. Passenger Demand Prediction with Cellular Footprints. In: *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Hong Kong, 2018, pp. 1-9. (In Eng.) DOI: 10.1109/SAHCN.2018.8397114
- [16] Nyhan M.M., Kloog I., Britter R., Ratti C., Koutrakis P. Quantifying population exposure to air pollution using individual mobility patterns inferred from mobile phone data. *Journal of Exposure Science & Environmental Epidemiology*. 2019; 29:238-247. (In Eng.) DOI: 10.1038/s41370-018-0038-9
- [17] Liu S. Navigation devices & usage - Statistics & Facts. Jan 31, 2018. Available at: <https://www.statista.com/topics/2221/navigation-devices-and-usage/> (accessed 16.05.2019). (In Eng.)
- [18] Zhang H., Shi B., Zhuge C., Wang W. Detecting Taxi Travel Patterns using GPS Trajectory Data: A Case Study of Beijing. *KSCE Journal of Civil Engineering*. 2019; 23(4):1797-1805. (In Eng.) DOI: 10.1007/s12205-019-0580-6
- [19] Hu X., An S., Wang J. Taxi Driver's Operation Behavior and Passengers' Demand Analysis Based on GPS Data. *Journal of Advanced Transportation*. 2018; 2018:6197549. (In Eng.) DOI: 10.1155/2018/6197549
- [20] Zhao J., et al. Truck Traffic Speed Prediction Under Non-Recurrent Congestion: Based on Optimized Deep Learning Algorithms and GPS Data. *IEEE Access*. 2019; 7:9116-9127. (In Eng.) DOI: 10.1109/ACCESS.2018.2890414
- [21] Galpern P., Ladle A., Uribe F.A., Sandalack B., Doyle-Baker P. Assessing urban connectivity using volunteered mobile phone GPS locations. *Applied Geography*. 2018; 93:37-46. (In Eng.) DOI: 10.1016/j.apgeog.2018.02.009
- [22] Zhai Yu., Baran P.K., Wu C. Can trail spatial attributes predict trail use level in urban forest park? An examination integrating GPS data and space syntax theory. *Urban Forestry & Urban Greening*. 2018; 29:171-182. (In Eng.) DOI: 10.1016/j.ufug.2017.10.008
- [23] Chen P., Shen Q., Childress S. A GPS data-based analysis of built environment influences on bicyclist route preferences. *International Journal of Sustainable Transportation*. 2018; 12(3):218-231. (In Eng.) DOI: 10.1080/15568318.2017.1349222
- [24] Deng M., Huang J., Zhang Yu., Liu H., Tang L., Tang J., Yang X. Generating urban road intersection models from low-frequency GPS trajectory data. *International Journal of Geographical Information Science*. 2018; 32(12):2337-2361. (In Eng.) DOI: 10.1080/13658816.2018.1510124
- [25] Yun H.J., Kang D.J., Lee M.J. Spatiotemporal distribution of urban walking tourists by season using GPS-based smartphone application. *Asia Pacific Journal of Tourism Research*. 2018; 23(11):1047-1061. (In Eng.) DOI: 10.1080/10941665.2018.1513949
- [26] Yang X., Stewart K., Tang L., Xie Z., Li Q. A Review of GPS Trajectories Classification Based on Transportation Mode. *Sensors*. 2018; 18(11):3741. (In Eng.) DOI: 10.3390/s18113741
- [27] Duan Z., Yang Y., Zhang K., Ni Y., Bajgain S. Improved Deep Hybrid Networks for Urban Traffic Flow Prediction Using Trajectory Data. *IEEE Access*. 2018; 6:31820-31827. (In Eng.) DOI: 10.1109/ACCESS.2018.2845863
- [28] Brezmes T., Gorricho J-L., Cotrina J. Activity Recognition from Accelerometer Data on a Mobile Phone. In: Omatu S. et al. (eds) *Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living. IWANN 2009. Lecture Notes in Computer Science*, vol. 5518. Springer, Berlin, Heidelberg, 2009, pp. 796-799. (In Eng.) DOI: 10.1007/978-3-642-02481-8_120
- [29] Zhao H., Hou C., Alrobassy H., Zeng X. Recognition of Transportation State by Smartphone Sensors Using Deep Bi-LSTM Neural Network. *Journal of Computer Networks and Communications*. 2019; 2019: 4967261. (In Eng.) DOI: 10.1155/2019/4967261
- [30] Ignatov A. Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *Applied Soft Computing*. 2018; 62:915-922. (In Eng.) DOI: 10.1016/j.asoc.2017.09.027
- [31] Almaslukh B., Al Muhtadi J., Artoli A. A robust convolutional neural network for online smartphone-based human activity recognition. *Journal of Intelligent & Fuzzy Systems*. 2018; 35(2): 1609-1620. (In Eng.) DOI: 10.3233/JIFS-169699
- [32] Lee K., Kwan M-P. Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results. *Computers, Environment and Urban Systems*. 2018; 67:124-131. (In Eng.) DOI: 10.1016/j.compenvurbsys.2017.09.012
- [33] Lu D-N., Nguyen D-N., Nguyen T-H., Nguyen H- Vehicle Mode and Driving Activity Detection Based on Analyzing Sensor Data of Smartphones. *Sensors*. 2018; 18(4):1036. (In Eng.)



- DOI: 10.3390/s18041036
- [34] Won M., Sayan S., Park K.-J. DeepWiTraffic: Low Cost Wi-Fi-Based Traffic Monitoring System Using Deep Learning. *CoRR*. 2018; abs/1812.08208. Available at: <http://arxiv.org/abs/1812.08208> (accessed 16.05.2019). (In Eng.)
- [35] Depatla S., Muralidharan A., Mostofi Y. Occupancy estimation using only WiFi power measurements. *IEEE Journal on Selected Areas in Communications*. 2015; 33(7):1381-1393. (In Eng.) DOI: 10.1109/jsac.2015.2430272
- [36] De Montjoye Y.-A., Hidalgo C.A., Verleysen M., Blondel V.D. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*. 2013; 3: 1376. (In Eng.) DOI: 10.1038/srep01376
- [37] Roth J.D., Tummala M., McEachen J.C., Scrofani J.W. Quantifying Location Privacy in Urban Next-Generation Cellular Networks. In: *Proceedings of the 51st Hawaii International Conference on System Sciences*. 2018, pp. 5831-5833. (In Eng.) DOI: 10.24251/hicss.2018.731
- [38] Smartphone data offered to exchange for wireless energy for charging. In: Indicator. 08.09.2018. Available at: <https://indicator.ru/mathematics/mobilnyj-kraudensing-08-09-2018.htm> (accessed 16.05.2019). (In Russ.)
- [39] Lane N.D., Chon Y., Zhou L., Zhang Y., Li F., Kim D., Ding G., Zhao F., Cha H. Piggyback CrowdSensing (PCS): energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities. In: *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys '13)*. Association for Computing Machinery, New York, NY, USA, 2013, Article 7, pp. 1-14. (In Eng.) DOI: 10.1145/2517351.2517372
- [40] Namiot D., Sneps-Snepp M. A Survey of Smart Cards Data Mining. In: *Supplementary Proceedings of the Sixth International Conference on Analysis of Images, Social Networks and Texts (AIST 2017)*. CEUR Workshop Proceedings. 2017; 1975:314-325. Available at: <http://ceur-ws.org/Vol-1975/paper33.pdf> (accessed 16.05.2019). (In Eng.)
- [41] Namiot D., Pokusaev O., Lazutkina V. On passenger flow data models for urban railways. *International Journal of Open Information Technologies*. 2018; 6(3):9-14. Available at: <https://elibrary.ru/item.asp?id=32595087> (accessed 16.05.2019). (In Russ., abstract in Eng.)
- [42] Namiot D.E., Pokusaev O.N., Kupriyanovsky V.P. On Patterns for the Use of Railway Stations. *Sovremennye informacionnye tehnologii i IT-obrazovanie = Modern Information Technologies and IT-Education*. 2018; 14(3):756-761. (In Russ., abstract in Eng.) DOI: 10.25559/SITI-TO.14.201803.756-761
- [43] Nekraplonna M., Namiot D. Metro correspondence matrix analysis. *International Journal of Open Information Technologies*. 2019; 7(7):68-80. Available at: <https://elibrary.ru/item.asp?id=38532115> (accessed 16.05.2019). (In Russ., abstract in Eng.)
- [44] Namiot D., Sneps-Snepp M. On Crowd Sensing Back-end. In: *Data Analytics and Management in Data Intensive Domains - Selected Papers (DAMDID/RCDL 2016)*. CEUR Workshop Proceedings. 2016; 1752:168-175. Available at: <http://ceur-ws.org/Vol-1752/paper28.pdf> (accessed 16.05.2019). (In Eng.)
- [45] Namiot D., Sneps-Snepp M. On Data Persistence Models for Mobile Crowdsensing Applications. In: Kalinichenko L., Kuznetsov S., Manolopoulos Y. (eds) *Data Analytics and Management in Data Intensive Domains. DAMDID/RCDL 2016. Communications in Computer and Information Science*, vol. 706. Springer, Cham, 2017, pp. 192-204. (In Eng.) DOI: 10.1007/978-3-319-57135-5_14
- [46] Daradkeh Y.I., ALdhaifallah M., Namiot D. Mobile Clouds for Smart Cities. *International Journal of Online and Biomedical Engineering*. 2017; 13(1):76-86. (In Eng.) DOI: 10.3991/ijoe.v13i01.6320
- Поступила 16.05.2019; принята к публикации 05.08.2019; опубликована онлайн 30.09.2019.
Submitted 16.05.2019; revised 05.08.2019; published online 30.09.2019.

Об авторах:

Булыгин Марк Валерьевич, магистрант факультета вычислительной математики и кибернетики, Московский государственный университет имени М. В. Ломоносова (119991, Россия, г. Москва, ГСП-1, Ленинские горы, д. 1), ORCID: <http://orcid.org/0000-0001-5846-4367>, messimm@yandex.ru

Намиот Дмитрий Евгеньевич, старший научный сотрудник лаборатории открытых информационных технологий, факультет вычислительной математики и кибернетики, Московский государственный университет имени М. В. Ломоносова (119991, Россия, г. Москва, ГСП-1, Ленинские горы, д. 1), кандидат физико-математических наук, ORCID: <http://orcid.org/0000-0002-4463-1678>, dnamiot@gmail.com

Все авторы прочитали и одобрили окончательный вариант рукописи.

About the authors:

Mark V. Bulygin, Master's Degree Student of the Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1, Leninskie gory, Moscow 119991, Russia), ORCID: <http://orcid.org/0000-0001-5846-4367>, messimm@yandex.ru

Dmitry E. Namiot, Senior Researcher of the Laboratory of Open Information Technologies, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1, Leninskie gory, Moscow 119991, Russia), Ph.D. (Phys.-Math.), ORCID: <http://orcid.org/0000-0002-4463-1678>, dnamiot@gmail.com

All authors have read and approved the final manuscript.

