

ПРОГНОЗИРОВАНИЕ ПАРАМЕТРОВ ДВИЖЕНИЯ ГОРОДСКОГО ПАССАЖИРСКОГО ТРАНСПОРТА ПО ДАННЫМ СПУТНИКОВОГО МОНИТОРИНГА

Агафонов А.А.^{1,2}, Сергеев А.В.^{1,2}, Чернов А.В.^{1,2,3}

¹ ОАО «Самара-Информспутник»,

² Институт систем обработки изображений РАН,

³ Самарский государственный аэрокосмический университет имени академика С.П. Королёва (национальный исследовательский университет)

Аннотация

Рассмотрены проблемы аналитики движения транспортных средств по данным спутникового мониторинга. Разработан алгоритм, обладающий преимуществами перед известными методами по точности и вычислительной эффективности. Составлена модель транспортной системы города. Разработана архитектура программного обеспечения, реализующего сервер аналитики. Проведены испытания точности прогнозирования прибытия на остановки муниципального транспорта г. Самары.

Ключевые слова: моделирование транспортной системы города, городской транспорт, аналитика движения, спутниковый мониторинг движения.

Введение

Возможность точного прогнозирования времени прибытия городского пассажирского транспорта на остановочные пункты является важным фактором развития транспортной системы города. Используя эту информацию, управляющая движением диспетчерская служба может быстро реагировать на непредвиденные дорожные ситуации, изменять маршруты движения транспорта, оптимизировать скорость и интенсивность движения транспортных средств.

Кроме того, прогноз прибытия транспорта и другие результаты аналитической обработки данных о движении транспорта должны быть доступны пассажирам – посредством сайта [1], мобильных приложений, СМС-сервисов, электронных информационных табло на остановках и других способов информирования.

В Самаре, как и в ряде других городов, оборудование ГЛОНАСС установлено на всём муниципальном транспорте. Кроме того, согласно новым правилам, которые вступили в силу с 1 июля 2012 года, компании, осуществляющие пассажирские перевозки, должны оборудовать подвижный состав транспорта ГЛОНАСС в рамках конкурсов на распределение маршрутов.

В данной работе рассматриваются преимущества и недостатки известных методов и описывается алгоритм работы аналитических сервисов программного комплекса Транспортного оператора Самары, разработанного авторами по заказу Администрации г.о. Самара. Алгоритм позволяет с высокой точностью и вычислительной эффективностью рассчитывать прогноз прибытия транспортных средств на остановочные пункты на основе данных спутникового мониторинга в реальном времени, модели транспортной системы города и планового времени прохождения маршрутов.

Обзор существующих методов

Существует несколько подходов к прогнозу времени прибытия транспортных средств: использование архивных данных и данных наблюдения в ре-

альном времени [2], использование статистических методов, использование техник машинного обучения (метода опорных векторов [3, 4], искусственных нейронных сетей), использование фильтрации (фильтр Калмана [4, 5]).

Подход с использованием архивных данных прогнозирует время прибытия транспортного средства в определённый период дня, основываясь на среднем времени прохождения участка дорожной сети в тот же период в предыдущие дни. Подход с использованием данных реального времени прогнозирует время прибытия транспорта, основываясь на положении и скорости транспорта в предыдущий момент времени. Недостаток этих подходов заключается в невозможности быстро реагировать на изменение дорожной ситуации и маршрутов движения.

Статистические модели прогнозируют время прохождения транспортом участка дорожной сети между остановками и затем определяют время прибытия на остановочные пункты. Эти модели обычно строятся как функции регрессии набора независимых переменных, таких как дорожные условия, пассажиропоток, погодные условия, задержки на остановках и т.д. Однако такой подход возможен лишь в случае, когда удаётся построить соответствующую функцию регрессии, что невозможно во многих прикладных областях, где переменные являются коррелированными. Кроме того, для предварительной настройки метода необходим значительный объём наблюдений, что трудно обеспечить в условиях большого города.

Методы машинного обучения позволяют учитывать сложные нелинейные соотношения между переменными, что делает их наиболее подходящими методами для прогнозирования времени прибытия. Применение дополнительной фильтрации позволяет объединить архивную информацию и данные реального времени для достижения более точного и надёжного прогноза. Однако эти методы требуют длительного обучения и являются вычислительно сложными, поэтому их трудно применять для большого количества транспортных средств.

Фильтр Калмана – рекурсивный фильтр, позволяющий получить более точную оценку прогноза, основываясь на исторических данных и данных реального времени. Однако методы, основанные на фильтрации Калмана, могут стать неэффективными при недоступности данных реального времени. Более того, на практике производительность фильтра Калмана сильно снижается из-за предположения, что шум процесса известен точно.

Исходные данные

Для прогнозирования времени прибытия транспортных средств на остановочные пункты доступны следующие данные из системы мониторинга подвижного состава:

- перечень транспортных средств, оснащённых системой мониторинга движения, выполняющих рейсы по маршрутам;
- координаты положений транспортных средств (в системе координат WGS-84) и моменты времени получения координат (x, y, t);
- модель транспортной системы города в виде ориентированного графа;
- плановое расписание прохождения транспортом контрольных точек маршрутов.

В получении данных существует ряд известных ограничений и недостатков:

- точность получаемых координат от датчика спутникового мониторинга составляет в среднем 6–8 метров, однако в некоторых случаях может достигать 70 метров;
- периодичность поступления информации о положении каждого транспортного средства в среднем составляет 30 секунд, строгой периодичности нет;
- возможны временные отключения датчиков и потери части информации о положении транспортных средств;
- необходимы проверки соответствия заявленного маршрута и направления фактически выполняемого движения (иногда данные содержат ошибки или актуализируются с опозданием);
- возможен уход транспортного средства со своего планового маршрута;
- расписания прохождения контрольных точек составлены без учёта зависимости от сезона, времени дня, погодных условий;
- не существует выполняемого норматива, какое время транспорт должен проводить на остановке. Количественные характеристики обрабатываемых данных следующие:
 - количество маршрутов, для которых ведётся мониторинг: 100–150 штук;
 - количество остановок, для которых рассчитываются прогнозы прибытия: 1000–1500 штук;
 - длина маршрута: 10–50 остановок, 3–40 километров;
 - количество транспортных средств, оснащённых системой мониторинга движения: 500–700 штук;
 - усреднённая периодичность обновления данных – 30 секунд;
 - количество запросов пользователей на получение аналитической и прогнозируемой информации – 5000 в минуту.

Дополнительные требования к алгоритмам обработки данных:

- учёт ограничения и недостатки в получении данных;
- вычисление прогнозов прибытия транспорта на остановки и других продуктов аналитической обработки с погрешностью не хуже 2 периодов обновления данных;
- время обработки по всем транспортным средствам и остановкам должно быть гарантированно меньше периода обновления данных («с запасом» в 2–3 раза).

Модель транспортной системы города

Важным этапом в предварительной подготовке исходных данных является построение модели транспортной системы города. Модель представляет собой ориентированный граф, по структуре и геометрии повторяющий правила прохождения маршрутов транспортом.

В силу пространственного характера модели логично использовать геоинформационную систему для подготовки исходных данных. В данной работе использовалась ГИС «Ингео» со специальным модулем, реализующим автоматические процедуры по подготовке транспортных данных. В качестве базовой карты использовалась Единая цифровая картографическая основа Самары в составе слоёв адресного плана, автомобильных дорог, трамвайных путей и путей метрополитена.

В нескольких слоях транспортной инфраструктуры была организована структура слоёв и стилей (классов и подклассов пространственных объектов) и семантических данных для описания остановок и маршрутов, а затем внесены данные по следующим правилам:

- остановочные пункты создавались как точечные объекты, расположенные на месте остановочных павильонов;
- маршруты автобусов, трамваев, троллейбусов и метро создавались в отдельных слоях по видам транспорта в виде объектов, содержащих в семантике упорядоченный список остановок в порядке их прохождения; два направления каждого маршрута создавались как два отдельных объекта;
- геометрическая форма маршрутов рассчитывалась автоматической процедурой по слоям осевых линий дорог и трамвайных путей, объединяя кратчайшими путями интервалы между последовательными остановками; результаты проверялись вручную, при необходимости корректировались промежуточными обязательными точками.

Объекты остановок и маршрутов также имели семантическое описание (название, направление) и классификационные номера, позволяющие однозначно устанавливать взаимосвязи между объектами.

Из подготовленных слоёв транспортной инфраструктуры автоматической процедурой строился следующий граф: остановки транспорта являлись вершинами графа, перегоны между остановками – рёбрами графа, причём каждая остановка представлялась как минимум тремя вершинами графа.

Такая структура графа удобна не только для определения позиции транспорта на маршруте и времени прибытия на остановки, но и для поиска маршрута проезда транспорта с учётом пересадок.

Часть графа, представляющего два маршрута R_1 и R_2 , содержащих общую остановку S_k , показана на рис. 1.

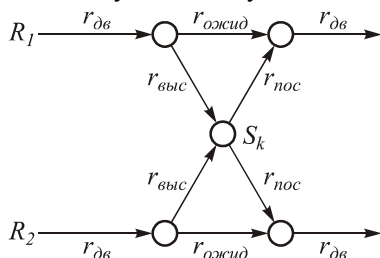


Рис. 1. Структура графа маршрутной сети

На рис. 1. $r_{дв}$ – ребро с весом – среднее время прохождения перегона между остановками; $r_{выс}$ – ребро с весом – среднее время высадки из транспорта; $r_{пос}$ – ребро с весами – среднее время посадки в транспорт и стоимость проезда; $r_{ожид}$ – ребро с весом – среднее времени ожидания транспорта на остановке.

Кроме того, остановки, расположенные в зоне пешеходной доступности, соединялись рёбрами с весами, равными среднему времени движения пешком.

К каждому ребру были приписаны три весовых характеристики: длина, время прохождения и стоимость прохождения ребра, которые формировались исходя из геометрии участка маршрута между остановками, проложенного по форме дорожной сети, трамвайных путей или линий метрополитена.

Среднее время посадки, высадки и задержки транспорта на остановках было получено по результатам наблюдений. Время прохождения перегона между остановками формировалось из планового расписания времени прохождения транспортом контрольных точек на маршруте.

Алгоритм работы сервера аналитики

Аналитическая и прогнозируемая информация формируется с помощью специального программного обеспечения – Сервера аналитики, который должен с высокой скоростью обрабатывать запросы на получение прогнозов и других продуктов аналитики, поступающие от сайтов, мобильных приложений, информационного табло и других потребителей. Общая архитектура программного обеспечения представлена на рис. 2.

Данные спутникового мониторинга поступают в сервер аналитики в режиме реального времени, и для поддержания актуальности результатов прогноза требуется постоянный пересчёт. Чтобы снизить зависимость вычислительной нагрузки сервера аналитики от количества поступающих запросов, используется программная архитектура из двух циклов обработки и хранилища промежуточных результатов расчёта в оперативной памяти. Первый цикл периодически запрашивает новые данные спутникового мониторинга, проводит расчёт прогнозов и со-

храняет промежуточные результаты. Второй цикл обрабатывает поступающие от потребителей запросы и отвечает на них по имеющимся промежуточным результатам. В таком варианте архитектуры нагрузка на сервер аналитики может быть очень высока, а актуальность предоставляемых прогнозов никогда не бывает хуже периода обновления данных.

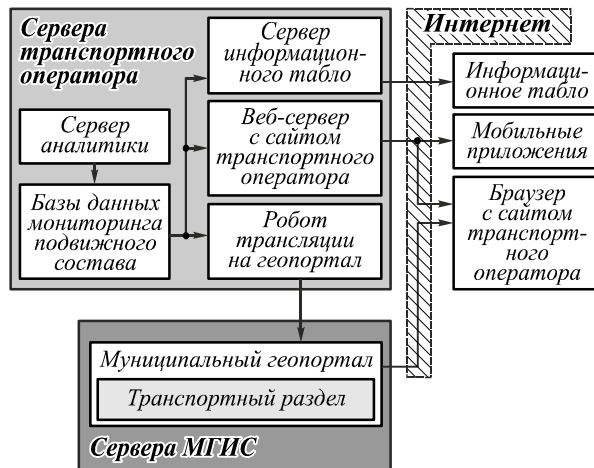


Рис. 2. Архитектура системы

Алгоритм циклического пересчёта состоит из следующих последовательных шагов:

- определяется полный набор транспортных средств, для которых ведётся спутниковый мониторинг, у базы данных мониторинга запрашиваются последние положения каждого из них;
- из набора выделяется подмножество транспортных средств, изменивших своё положение с предыдущего цикла пересчёта, для каждого них выполняются последующие шаги;
- по модели транспортной системы города делается уточнение положения транспорта, в какой координате и на каком участке маршрута находится транспорт;
- по положению транспорта определяется скорость движения (из свойств рёбер графа), уточняется направление, принимается решение о сходе с маршрута при значительном отклонении;
- по положению транспорта для ближайшей остановки по ходу следования остановки рассчитывается время, которое требуется транспорту до её достижения, а для всех последующих оценка времени достижения делается накоплением времени прохождения рёбер графа по цепочке маршрута;
- в хранилище промежуточных результатов заносится новая порция прогнозов от рассматриваемого транспорта.

Важным свойством алгоритма является то, что расчёт требует значений времени прохождения только тех немногих участков маршрута, которые соответствуют текущему положению транспортного средства, появляется возможность рассчитывать эти характеристики моментально, не теряя в вычислительной эффективности всей системы. Значения времени прохождения участка маршрута вычисляются из известных фактов прохождения участка за последние минуты, часы, дни и месяцы, а также времени прохождения маршрутов, назначаемого при

планировании. Алгоритму предоставляется возможность использовать сразу всю эту информацию для каждого отдельного участка маршрута.

Благодаря этому удаётся избежать недостатков других существующих методов прогнозирования. В отличие от методов с использованием архивных данных и данных наблюдения в реальном времени, алгоритм может реагировать на резкие изменения дорожной ситуации. В отличие от статистических методов, техник машинного обучения и фильтрации Калмана, алгоритм способен работать без большого количества обучающих данных и не терять вычислительной эффективности при нехватке данных для отдельных участков маршрутной сети.

Алгоритм получения прогнозов для заданной остановки состоит из двух шагов:

- из хранилища делается выборка прогнозов всех маршрутов, проходящих через остановку;
- набор прогнозов сортируется по времени, снабжается дополнительной информацией о прибывающих транспортных средствах (бортовой номер, местонахождение) и в определённом формате записи выдаётся в ответ на запрос.

Оценка качества прогнозирования

Испытания разработанного сервера аналитики проводились в Самаре на нескольких наиболее нагруженных остановках – трамвайной «КРЦ Звезда» на улице Ново-Садовой и автобусно-троллейбусной «Аэрокосмический университет» на Московском шоссе. Исследования были проведены в часы пик (8:00–10:00) и в часы свободного движения (13:00–15:00). Сравнивались результаты прогнозов и моменты фактического прибытия транспорта на остановки.

Графики соответствия прогнозируемого и фактического времени прибытия транспортных средств в часы свободного движения представлены на рис. 3 (автобусы), 4 (троллейбусы) и 5 (трамваи). Для каждой точки горизонтальное положение означает полученный в некоторый момент времени прогноз, а вертикальное положение – действительное время, спустя которое транспорт прибыл на остановку. Точки, находящиеся на диагонали, соответствуют безошибочным прогнозам.

Графики соответствия прогнозируемого и фактического времени прибытия транспортных средств в часы пик представлены на рис. 6 (автобусы), 7 (троллейбусы) и 8 (трамваи).

По графикам заметны общие особенности прогнозов: в часы свободного движения транспорт, как правило, прибывает раньше, а в часы пик – позже прогнозируемого времени.

Эти недостатки точности прогнозирования могут быть устранены, если варьировать в течение суток время прохождения каждого ребра в модели транспортной системы города, используя архивные данные о движении и статистический метод, а также оперативную информацию («пробки») о средних скоростях движения на участках транспортной сети.

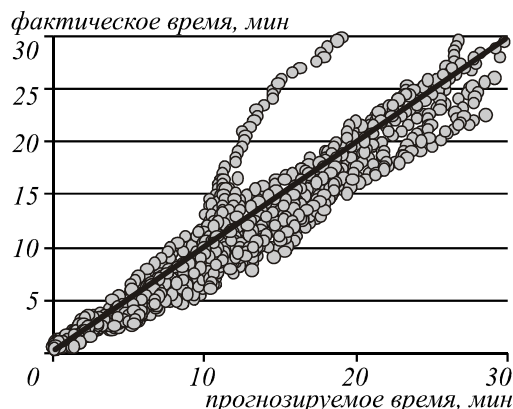


Рис. 3. График для автобусов при свободном движении

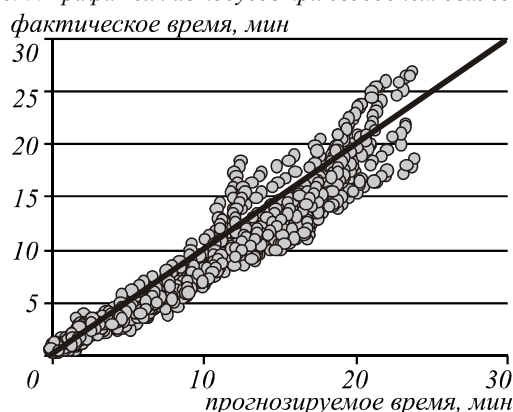


Рис. 4. График для троллейбусов при свободном движении

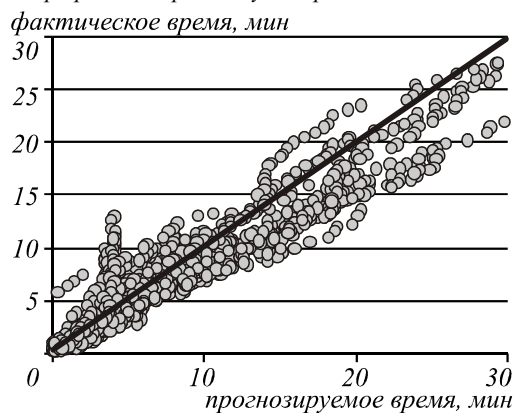


Рис. 5. График для трамваев при свободном движении

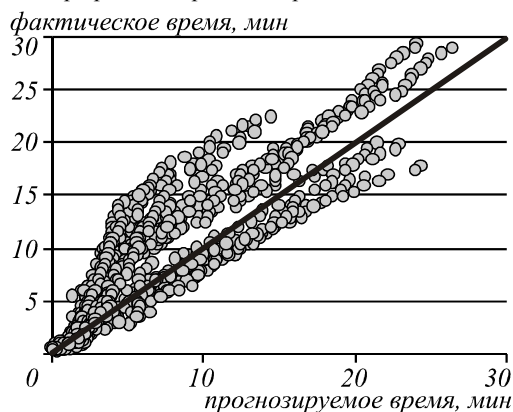


Рис. 6. График для автобусов в часы пик

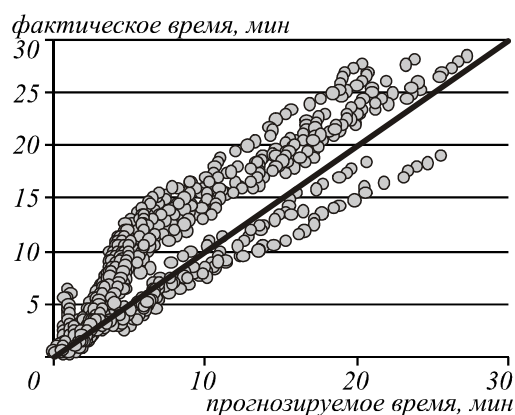


Рис. 7. График для троллейбусов в часы пик

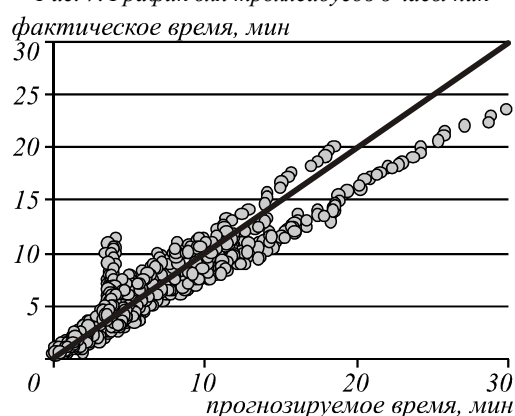


Рис. 8. График для трамваев в часы пик

Заключение

В статье представлена архитектура системы предоставления пассажирам информации о движении транспорта. Система обрабатывает в реальном времени 1345 транспортных средств, автоматически определяет направление маршрута и прогнозирует время прибытия на остановочные пункты. Система поддерживает оповещение пассажиров через сайт [1], мобильные приложения и электронные табло на остановках.

Предложен алгоритм прогнозирования времени прибытия транспортных средств на остановочные пункты, использующий плановую скорость движения на участках маршрутной сети.

Данный алгоритм не позволяет быстро реагировать на изменение дорожной ситуации. Для улучшения прогнозов необходимо использовать архивные данные с учётом данных реального времени.

Благодарности

Работа выполнена при поддержке РФФИ (гранты 11-07-12059-офи-м, 11-07-12060-офи-м).

Литература

1. Сайт транспортного оператора Самары <http://tosamara.ru/>.
2. **Pu, W.** Real-Time Estimation of Urban Street Segment Travel Time Using Buses as Speed Probes / W. Pu, J. Lin and L. Long // Transportation Research Record: Journal of the Transportation Research Board. – 2009. – Vol. 2129. – P. 81-89.
3. **Ma, J.** Accurate On-line Support Vector Regression / Junshui Ma, James Theiler, Simon Perkins // Journal: Neural Computation – NECO. – 2003. – Vol. 15, N 11. – P. 2683-2703.
4. **Ma, Zh.** An Aggregation Method for Dynamic Bus Arrival Time Prediction / Zhenliang Ma, Xing Jianping, Yan Shihao, Wu Yong, Wu Yubing // AISS. – 2011. – Vol. 3, N 6. – P. 37-48.
5. **Chu, L.** Adaptive kalman filter based freeway travel time estimation / L. Chu, J.S. Oh, W. Recker // In 84th TRB Annual Meeting, Washington, 2005.

References

1. Samara Transport Operator official site <http://tosamara.ru/en/>.
2. **Pu, W.** Real-Time Estimation of Urban Street Segment Travel Time Using Buses as Speed Probes / W. Pu, J. Lin and L. Long // Transportation Research Record: Journal of the Transportation Research Board. – 2009. – Vol. 2129. – P. 81-89.
3. **Ma, J.** Accurate On-line Support Vector Regression / Junshui Ma, James Theiler, Simon Perkins // Journal: Neural Computation – NECO. – 2003. – Vol. 15, N 11. – P. 2683-2703.
4. **Ma, Zh.** An Aggregation Method for Dynamic Bus Arrival Time Prediction / Zhenliang Ma, Xing Jianping, Yan Shihao, Wu Yong, Wu Yubing // AISS. – 2011. – Vol. 3, N 6. – P. 37-48.
5. **Chu, L.** Adaptive kalman filter based freeway travel time estimation / L. Chu, J.S. Oh, W. Recker // In 84th TRB Annual Meeting, Washington, 2005.

FORECASTING OF THE MOTION PARAMETERS OF CITY TRANSPORT BY SATELLITE MONITORING DATA

A.A. Agafonov^{1,2}, A.V. Sergeyev^{1,2}, A.V. Chernov^{1,2,3}

¹ OJSC «Samara-Informsputnik»,

² Image Processing Systems Institute of the RAS,

³ S.P. Korolyov Samara State Aerospace University (National Research University)

Abstract

Problems of analytics of vehicles movements according to satellite monitoring data are considered. The algorithm with the better accuracy and computing efficiency in comparison with known methods is developed. The model of the transport system of the city is created. The architecture of the software realizing the analytics server is developed. The method for arrival time prediction is examined with the data of municipal transport in Samara, Russia.

Key words: modeling of the transport system of the city, municipal transport, analytics of vehicles movement, satellite monitoring of movement.

Сведения об авторах

Агафонов Антон Александрович, 1988 года рождения. В 2011 году окончил Самарский государственный аэрокосмический университет (СГАУ). В настоящее время работает стажёром-исследователем в Институте систем обработки изображений РАН и, по совместительству, инженером-математиком в ОАО «Самара-Информспутник». Круг научных интересов включает геоинформационные технологии, веб-технологии. Имеет 1 публикацию.

E-mail: ant.agafonov@gmail.com .

Anton Aleksandrovich Agafonov (1988 b.), graduated from S.P. Korolyov Samara State Aerospace University (SSAU) at 2011. At present he is intern-researcher at the Image Processing Systems Institute of the Russian Academy of Sciences, holding a part-time positions of engineer-mathematician at JSC "Samara-Informsputnik". The area of interests includes geoinformatics and web-technologies. He's list of publications contains 1 paper.

Сергеев Александр Владиславович, 1986 года рождения. В 2009 г. окончил Самарский государственный аэрокосмический университет (СГАУ). В настоящее время работает стажёром-исследователем в Институте систем обработки изображений РАН и, по совместительству, инженером-математиком в ОАО «Самара-Информспутник». Круг научных интересов включает обработку изображений, геоинформационные технологии, веб-технологии. Имеет 7 публикаций, из них 1 статью. Является членом Российской ассоциации распознавания образов и анализа изображений.

E-mail: a.sergeyev@gmail.com .

Aleksandr Vladislavovich Sergeev (1986 b.), graduated from the S.P. Korolyov Samara State Aerospace University (SSAU) at 2009. At present he is intern-researcher at the Image Processing Systems Institute of the Russian Academy of Sciences, holding a part-time position of engineer-mathematician at JSC "Samara-Informsputnik". The area of interests includes digital signals and image processing, geoinformatics and web-technologies. He's list of publications contains 7 papers, including 1 article. He is a member of Russian Association of Pattern Recognition and Image Analysis.



Чернов Андрей Владимирович, 1975 года рождения. В 1998 году окончил Самарский государственный аэрокосмический университет (СГАУ). В 1998 году поступил в аспирантуру СГАУ, в 2004 году защитил диссертацию на соискание степени кандидата технических наук. В настоящее время работает старшим научным сотрудником в Институте систем обработки изображений РАН и, одновременно, доцентом кафедры геоинформатики и информационной безопасности СГАУ. Круг научных интересов включает цифровую обработку сигналов и изображений, геоинформатику, обработку данных дистанционного зондирования. Имеет более 50 публикаций, в том числе 15 статей и две монографии (в соавторстве). Член Российской ассоциации распознавания образов и анализа изображений.

E-mail: aches@smr.ru .

Andrey Vladimirovich Chernov (1975 b.), graduated (1998) from the S.P. Korolyov Samara State Aerospace University (SSAU). He received his PhD in Technical sciences (2005). At present he is a researcher at the Image Processing Systems Institute of the Russian Academy of Sciences, and, holding a part-time position of Associate Professor at SSAU's Geoinformatics and Information Security sub-department. The area of interests includes digital signals and image processing, geoinformatics, remote sensing. He's list of publications contains more than 50 scientific papers, including 15 articles and 2 monographs. He is a member of Russian Association of Pattern Recognition and Image Analysis.

Поступила в редакцию 17 июня 2012 г.