

## **Краткосрочное прогнозирование скоростей движения автомобильного транспорта с использованием алгоритмической композиции**

Н.В. Антипова<sup>1</sup>, Ю.В. Чехович<sup>2</sup>

1 – Московский физико-технический институт (государственный университет)

2 – Вычислительный центр им. Дороницына РАН

В крупных городах движение автотранспорта часто затруднено, водителей волнуют два вопроса: сколько времени будет потрачено на дорогу, и какой маршрут выбрать. Второй вопрос особенно важен, средняя продолжительность маршрута 1-2 часа, ситуация за это время может измениться, нужно прогнозирование на 1-2 часа. Для ответа на эти вопросы существуют специальные интернет-сервисы, один из них - Яндекс.Пробки. Сервис предлагает прогноз пробок, однако использует прогнозирование константой, что не даёт точных результатов в силу динамичности ситуации на дороге. Таким образом, есть потребность в создании алгоритма прогноза высокой точности.

По данной теме были проведены исследования по прогнозированию загруженности дорог различными алгоритмами. По итогам исследований был создан алгоритм прогнозирования, который показал результаты лучше, чем константный. Данная работа является продолжением аналогичной работы тех же авторов и промежуточной ступенью к завершающему этапу исследований: проверке введённых алгоритмов на новых данных.

Цель данной работы - проверить работу введённого алгоритма прогнозирования и зависимость качества его работы от введённых параметров.

Работа проводилась на данных, содержащих описание транспортных рёбер, их топологию, а также значения скоростей каждые 2 минуты. В качестве информации об ошибках прогнозирования различными алгоритмами использовался трёхмерный массив, где каждой тройке (ребро, момент времени, зона горизонта прогноза) ставилась в соответствие ошибка прогноза для каждого из рассмотренных алгоритмов. Данные об ошибках имеются за три недели. Обучение проводилось на первой и второй неделях, контроль - на третьей неделе.

Задача классификации решалась при помощи построения деревьев классификации на основе признакового описания объектов и значений ошибок прогнозирования. После построения дерева проводилась первая проверка корректности, при необходимости проводилось дообучение, затем проводились вычисления на контрольной выборке.

Для введения типизации объектов классификации было создано признаковое описание: были учтены различные значения скоростей, насыщенности данными, топологические признаки, день недели и время суток. По сравнению с предыдущей работой количество признаков в признаковом описании было увеличено с 13 до 40, также рассмотрена задача двухклассовой классификации и не рассмотрена задача многоклассовой ввиду результатов предыдущего исследования.

Собственно, сама задача двухклассовой классификации заключается в поочерёдном сравнении алгоритмов: константного алгоритма с алгоритмом, дающим самую маленькую ошибку на всей неделе, при успешной работе классификатора на первом контроле в рассмотрение добавлялся следующий по величине суммарной за неделю ошибке алгоритм, и так до тех пор, пока не будет достигнуто оптимальное по величине полученной ошибки количество алгоритмов в композиции.

В задаче классификации на этапе обучения для различных зон прогноза были подобраны различные количества алгоритмов для композиции. Проверка на контрольной выборке показала, что данный алгоритм классификации является оптимальным для решения задачи прогнозирования в сравнении с прогнозированием константой, введение дополнительных признаков также уменьшило погрешность прогнозирования.

В данной работе было составлено расширенное признаковое описание объектов, была решена задача двухклассовой классификации, в результате чего удалось убедиться в корректности выбранного универсального алгоритма.

### Литература

1. Автомобильное издание Lenta.ru: количество машин на дорогах Москвы [Электронный ресурс] – Режим доступа: <http://motor.ru/news/2010/11/18/moscow/>, свободный. – Загл. с экрана.
2. Мобильный Яндекс.Карты: схема более 380 городов для мобильных устройств [Ссылка на приложение] – Режим доступа: <http://mobile.yandex.ru/apps/maps/>, свободный. – Загл. с экрана.
3. Геоинформационные технологии: как работают Яндекс.Пробки [Электронный ресурс] – Режим доступа: <http://company.yandex.ru/technologies/yaprobki/>, свободный. – Загл. с экрана.
4. Хохлов М.А. Математические модели, использующиеся в Яндекс.Пробках: методы прогнозирования транспортной ситуации [Видеозапись конференции] / Яндекс.Технологии; Yet another Conference 2012; – М., [2012]. – Режим доступа: <http://tech.yandex.ru/events/yac/2012/talks/382/>, свободный. – Загл. с экрана.
5. Ивкин Н.П., Чехович Ю.В. Краткосрочное прогнозирование скоростей транспортных потоков // Интеллектуализация обработки информации: 9-я международная конференция. Республика Черногория, г. Будва, 2012 г.: Сборник докладов. – М.: Торус Пресс, 2012. С. 219-222.
6. Центр компетенций MathWorks: Функции нелинейного регрессионного анализа на основе графа возможных решений [Электронный ресурс] – Режим доступа: <http://matlab.exponenta.ru/statist/book2/17/treefit.php>, свободный. – Загл. с экрана.
7. Антипова Н.В., Чехович Ю.В. Краткосрочное прогнозирование скоростей движения автомобильного транспорта с использованием алгоритмической композиции // «Труды Московского физико-технического института» - 2014.