

УДК 656.11.021.2

С. Г. Селевич, канд. техн. наук, доц. НТУ «ХПИ»**МЕТОДИКА ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ДАННЫХ
АВТОМОБИЛЬНЫХ GPS-НАВИГАТОРОВ ДЛЯ РЕШЕНИЯ ЗАДАЧ
ПРОГНОЗИРОВАНИЯ ТРАНСПОРТНЫХ ПОТОКОВ**

Предложена методика кластеризации данных автомобильных GPS-навигаторов для последующего построения прогнозных моделей транспортных потоков.

Ключевые слова: GPS-навигатор, транспортный поток, прогнозная модель, кластеризация.

Введение. В среднем по Украине на июль 2008 года показатель автомобилизации составлял 157 авто на 1000 жителей и несмотря на кризис постепенно увеличивается. При этом на относительно ограниченной территории (2-5%), в основном крупных городов, сосредоточено от 50 до 90% населения и около 45% парка легковых автомобилей [1]. К числу характерных проблем, с которыми сталкивались все мировые мегаполисы при достижении критического уровня автомобилизации, относятся: резкое увеличение уровня загрязнения окружающей среды и шумов, чрезмерный расход топлива вследствие снижения средней скорости движения транспортных потоков до 15-30км/ч в период пиковых нагрузок, высокий уровень аварийности [2]. Так, величина издержек в странах Европейского союза в процентах от ВВП составляет: от перегруженных транспортных инфраструктур – 2,0%, последствий аварий на транспорте – 1,5%, от негативного воздействия транспорта на окружающую среду – 0,6%.

Мировой опыт борьбы с последствиями урбанизации и ростом автомобилизации связан с внедрением совокупности мероприятий, среди которых можно выделить основные направления: искусственное ограничение количества легковых автомобилей на дорогах на законодательно-государственном уровне, реконструкция дорожно-транспортной инфраструктуры и использование интеллектуальных методов организации дорожного движения. Возможности первого и второго подхода в значительной мере исчерпаны, в то время как, эффективность внедрения интеллектуальных систем регулирования дорожного движения во многом обусловлена степенью адекватности математических моделей транспортных потоков заложенных в них.

Анализ публикаций. Наиболее полная классификация подходов к построению математических моделей транспортных потоков приведена в [2,3]. При этом базисом для их параметрического синтеза или проверки адекватности являются измерения параметров транспортного потока в реальном режиме времени. Кроме того, непрерывный мониторинг транспортных потоков позволяет оценить эффективность организации системы регулирования дорожного движения на протяжении длительных временных интервалов и в процессе внезапных изменений.

На практике для сбора данных используются преимущественно локальные детекторы [4,5], позволяющие измерять скорости транспортного потока на конкретном участке дорожной сети. Основным ограничением на пути внедрения таких систем, в том числе и в Украине, является их стоимость. Альтернативным подходом является

© С. Г. Селевич, 2012

использование автомобильных GPS навигаторов [6] в том числе встроенных в мобильные телефоны водителей [7], способных посредством существующих каналов GSM/CDMA каналов связи передавать свои координаты (треки) и скорость движения с достаточной степенью точности.

Благодаря достаточно высокой степени проникновения мобильной связи (в Украине по данным GfK Ukraine на конец 2010г составил 82%) такой подход к сбору информации о транспортных потоках представляется наиболее перспективным. При этом помимо задачи привязки треков к местности, остро становится вопрос обработки базы данных скоростей движения.

Цель и постановка задачи. Целью данной работы является обоснование методики предварительной обработки массива скоростей, формируемого автомобильными GPS-навигаторами, для их использования при построении прогнозных моделей транспортных потоков мегаполиса. В качестве исходных данных для анализа были использованы материалы [8].

Особенности данных о скоростях транспортного потока автомобильных GPS-навигаторов. Одним из основных ограничений в использовании данных автомобильных GPS-навигаторов является их неоднородность, которая заключается в том, что для отдельных участков дорожной сети известно относительно большое число значений скоростей, в то время как для участков с малой концентрацией автомобилей – недостаточное для анализа и формулирования выводов. На рис. 1, в качестве примера, представлено распределение количества данных о скоростях движения транспортных средств по участкам дорожной сети.

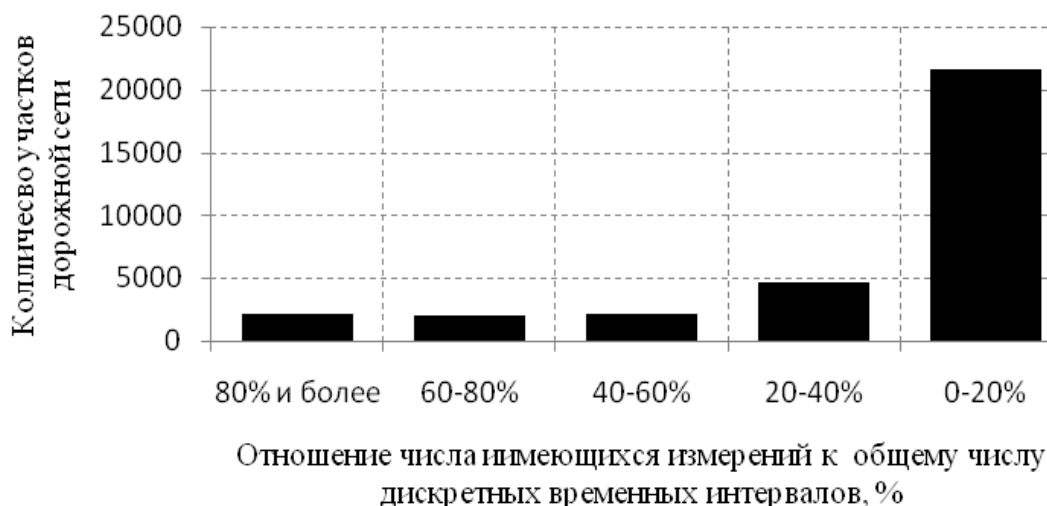


Рисунок 1 – Гистограмма наличия данных по участкам дорожной сети

Наличие малого количества измерений характерно для менее загруженных дорог, как правило, находящихся за пределами «бизнес-части» мегаполиса, для которых задача построения модели и прогнозирования скорости транспортного потока стоит менее остро по отношению к центральным частям. На участках с числом измерений $n > 300$ доверительный интервал на уровне значимости $q=0,1$ для математического ожидания скорости транспортного потока составляет [19, 96] км/ч, а

для среднеквадратического изменения - [6, 36] км/ч. Принимая также во внимание вероятно высокий уровень ошибки прогнозирования, в дальнейшем будем рассматривать лишь те участки, для которых известно хотя-бы по одному измерению для более чем 10% временных интервалов.

Неоднородность данных проявляется также и в том, что на отдельных, даже малых, временных интервалах сконцентрировано несколько измерений, а для других – наблюдается их отсутствие, как показано на рис. 2. При этом с точки зрения построения прогнозных моделей наиболее неблагоприятными являются длительные «пропуски» по временной оси, для которых точность интерполяции резко снижается.

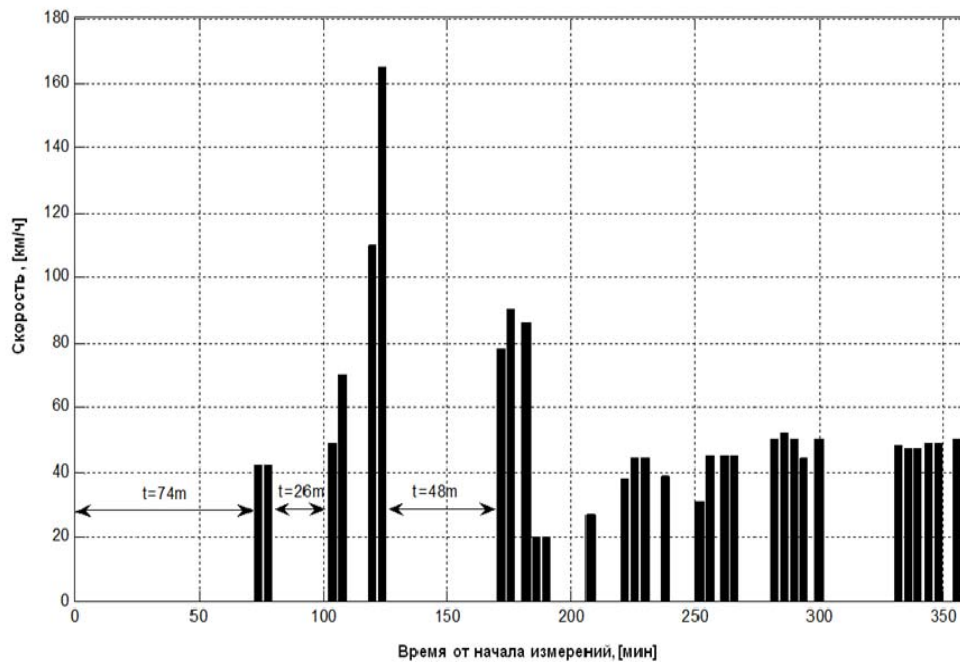


Рисунок 2 – Распределение измерений на одном участке дороги по времени опыта

Одной из общих особенностей, находящей подтверждение при анализе данных автомобильных GPS-приемников является отличие средних скоростей и их среднеквадратических отклонений в зависимости от характера дня недели: выходной или рабочий. При этом случайные колебания скорости не дают оснований для выделения других закономерностей, связанных с учетом дня недели в общем случае.

Кластеризация массива транспортных потоков. Представляется целесообразным на этапе предварительной обработки все участки разделить на группы в зависимости от величины среднеквадратического отклонения (СКО) скорости на участке, что позволит отдельно обрабатывать участки с относительно постоянной скоростью и те участки, где отклонение достаточно велико, а также в зависимости от времени суток, что позволит учесть влияние так называемого «часа пик» на среднюю скорость движения.

В ходе кластеризации по величине СКО (табл.) числовые значения границ интервалов получены по критерию минимума внутрикластерной суммы расстояний точек кластера до его центра. Так для будних дней недели получены более широкие (на 1 км/ч) интервалы второго кластера, что говорит о большем разбросе результатов наблюдений относительного среднего по отношению к выходным дням.

Таблица – Результаты кластеризации СКО скорости транспортных потоков

Номер кластера		I	II	III
Середина кластера, км/ч	выходные дни	12	18	25
	будние дни	9	16	28
Границы кластера, км/ч	выходные дни	[0;15)	[15;21)	(21;∞)
	будние дни	[0;13)	[13;22)	(22;∞)

Характер изменения средней скорости движения (см. рис. 3) на исследуемом интервале [16:00,22:00] явно свидетельствует о её росте с течением времени вне зависимости от типа дня недели (выходной/будний). Указанная особенность лишь подтверждает ожидаемую тенденцию освобождения автомобильных дорог по мере отдаления от «часа пик» и будет учтена в дальнейшем при построении прогнозных моделей. При этом разделение на интервалы с резко отличающимися характеристиками целесообразно проводить лишь для будних дней, для которых в общем случае характерно наличие точки «перелома», где снижение средней скорости с течением времени сменяется ростом.

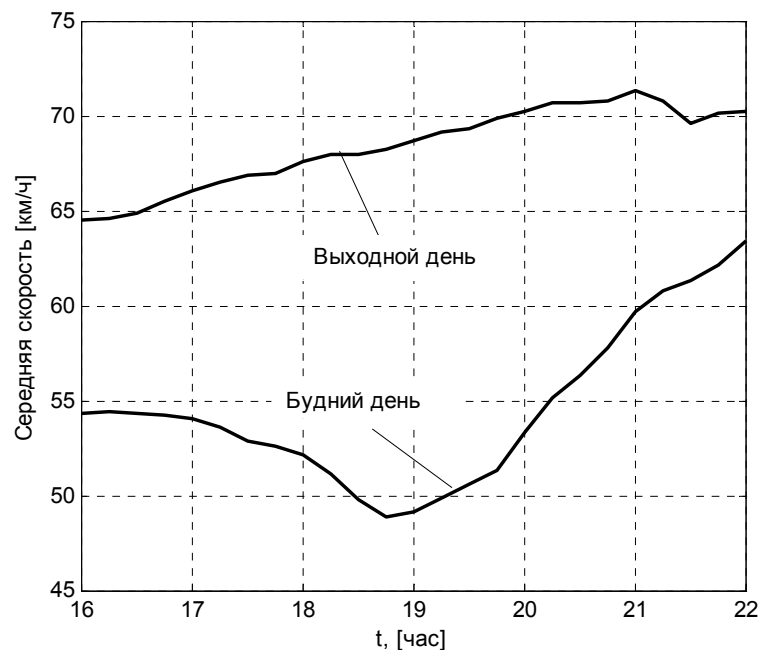


Рисунок 3 – Зависимость средней скорости движения транспортных потоков от времени суток.

Для большинства дорог, вне зависимости от их местоположения, перелом характеристики средней скорости движения находится в окрестности 18:45 и несколько смещен в сторону 19:00. С учетом того, что на этапе построения прогнозных моделей для удобства компьютерной обработки данные планируется приводить к равноотстоящим (15-ти минутным) интервалам, весь исследуемый участок времени для будних дней разделен на интервалы времени: [16:00,18:45) и [18:45, 22:00].

Выводы

Одним из наиболее эффективных подходов к получению достоверных характеристик транспортных потоков является использование данных автомобильных GPS-навигаторов. При этом выполнять построение прогнозных моделей, основанных на численных характеристиках транспортных потоков, целесообразно лишь для участков с достаточным количеством измерений и отдельно с учетом предложенной кластеризации: для выходных и рабочих дней, для времени суток и групп участков со схожими среднеквадратическими отклонениями скорости.

Список литературы: 1. Дубова С.В., Васильева А.Ю., Сильчук В.А. //Методы ограничения легкового транспорта в городах. – Містобудування та територіальне планування. Науково-технічний збірник КНУБА, випуск №32. – Київ, КНУБА, 2009. 2. Семенов В.В. Математические методы моделирования транспортных потоков // Сборник «Новое в синергетике. Новая реальность, новые проблемы, новое поколение». М: Наука, 2007. С. 102–133. 3. Швецов В.И. Математическое моделирование транспортных потоков // Автоматика и телемеханика. – 2003. – №11. 4. S. Lämmer and D. Helbing (2010) Self-Stabilizing Decentralized Signal Control of Realistic, Saturated Network Traffic. Santa Fe Working Paper Nr. 10-09-019. 5. «Видеопробки» [Электронный ресурс] // <http://videoprobki.ua/>. 6. Богумил В.Н., Ефименко Д.Б. Экспериментальные исследования транспортных потоков с использованием навигационных данных (глонасс/gps) диспетчерских систем // Наука и техника в дорожной отрасли. М. 2011. № 4. с. 3а-7. 7. J. C. Herrera, et al., “Evaluation of Traffic Data Obtained via GPS-Enabled Mobile Phones: The Mobile Century Field Experiment,” Transportation Research Part C, Vol. 18, No. 4, 2010, pp. 568-583. 8. «Интернет-математика - 2010» [Электронный ресурс] // «Яндекс». DOI= <http://imat2010.yandex.ru/>.

Поступила в редколлегию 15.11.2012

УДК 656.11.021.2

Методика редварительной обработки данных автомобильных GPS-навигаторов для решения задач прогнозирования транспортных потоков / С.Г. Селевич // Вісник НТУ «ХПІ». Серія: Автомобіле- та тракторобудування, 2012. – № 60 (966). – С. 92–96. – Бібліогр.: 8 назв.

Запропонована методика кластеризації даних автомобільних GPS-навігаторів для подальшої побудови прогнозних моделей транспортних потоків.

Ключові слова: GPS-навігатор, транспортний потік, прогнозна модель, кластеризація.

The methodology of automobile GPS-navigators data clustering is proposed for the transport flows prediction models construction.

Keywords: GPS-navigator, traffic flow, forecasting models, clustering.