

Виртуальный детектор транспорта

Исследуется задача о возможности и ограничениях определения интенсивности транспортного потока, на основании данных GPS-треков и информации о светофорном регулировании во время проезда транспортных средств с трекерами на борту с помощью имитационной модели движения транспортного потока.

1. Введение

Измерение интенсивности транспортного потока (ТП), достаточно сложная и ресурсоемкая задача. Стационарный детектор транспорта (ДТ), любого типа (инфракрасный, электромагнитный, ультразвуковой, видеофиксация) требует наличия самого устройства, канала связи, электропитания и оборудования. Также, требуется постоянная подстройка поля обзора и корректировка с учетом сезона. Один из главных недостатков - это то, что один детектор может обслуживать постоянно только один перекресток (точку на магистрали) или даже только одну полосу, а для построения различных моделей и управления, нужна по возможности наиболее полная информация о ТП в районе управления.

В связи с широким распространением устройств GPS навигации и смартфонов, появилась надежда на то, что учет характеристик ТП можно будет вести по каждому участнику индивидуально. Однако, такое возможно только в случае, если все или значительное большинство участников движения будет сообщать условному "центру" о своем перемещении, что в настоящее время кажется невозможным. Учитывая групповой характер движения, по тем данным, что имеются, можно судить о средней скорости потока, но определить точно значение интенсивности не представляется возможным.

2. Виртуальный детектор транспорта

Виртуальный детектор транспорта (ВДТ) - комплекс программ (алгоритмов) позволяющих восстановить значения интенсивности ТП в определенный момент времени в некоторых точках дорожной сети, например стоп линиях перекрестков, с приемлемой точностью, на базе модели движения ТП и информации о:

1. графе дорожной сети;
2. программах светофорного регулирования на перекрестках;
3. множестве GPS треків в координатах (lat, long, time).

На рисунке 1 представлена схема минимального участка дорожной сети и граф потоков, для которого имеется возможность определить интенсивность потоков на дугах DE, в первую очередь и AD, BD, CD во вторую. A,B,C - стоп линии первого перекрестка, D - точка выхода из перекрестка №1 в направлении перекрестка №2, E - стоп линия второго перекрестка.

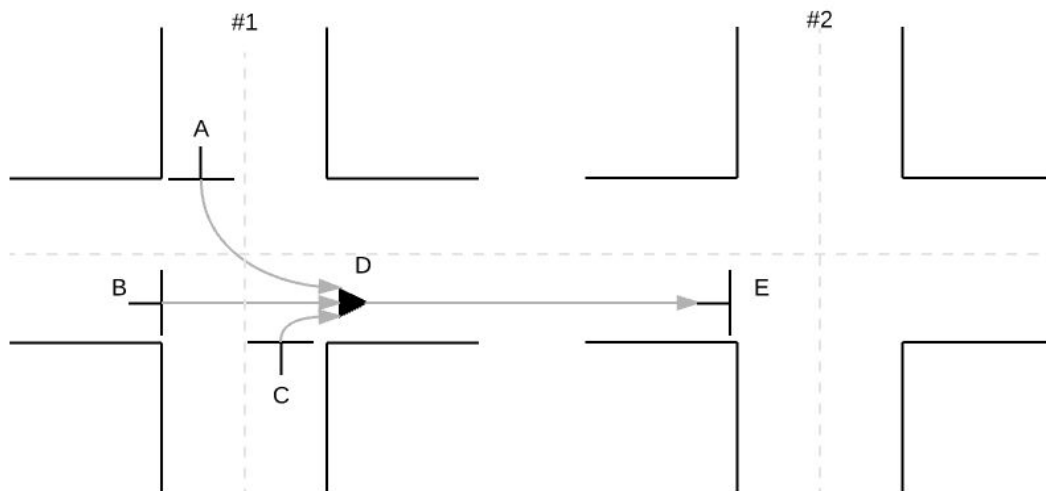


Рис 1. Схема потоков на 2-х смежных перекрестках

Идея о возможности построения ВДТ базируется на выполнении одного условия и верности двух предположений.

Условие 1. Известны длины всех дуг (AD, BD, CD, DE), скорость свободного проезда по ним, пропускные способности, кол-во полос на проезжей части в каждой точке, программы светофорного регулирования на обоих перекрестках - длительности сигналов/циклов смещение программ относительно друг-друга.

Предположение 1. Для восстановления значения интенсивности в точке E необходимо, чтобы в данный момент времени (например за последние 10 мин) иметь хотя бы один измеренный GPS-трек ТС, которое проехало через два смежных перекрестка, и остановилось на втором из них на запрещающий сигнал светофора.

Предположение 2. Для более точного определения интенсивности, необходимо знать отношение количества ТС прошедших через точки A, B, C, за указанные 10 мин, к общей сумме ТС в точке D ($\frac{Q_{AD}}{Q_D}, \frac{Q_{BD}}{Q_D}, \frac{Q_{CD}}{Q_D}$). Будем считать, что в точку E приезжают все ТС из D.

Имея достаточное количество треков, которые проезжали через точки (A, B, C) перекрестка №1, за продолжительный отрезок времени, например год или два. Можно провести кластеризацию этих треков, в первом приближении просто по календарю. То есть будут сгруппированы все треки в то же время суток 10 минут(час) в похожие дни (сезон, рабочий или выходной день и т.д.).

Суть предположения в том, что распределение вероятностей разъезда/съезда на входах/выходах перекрестка в пределах одного временного кластера должно быть неизменным и его можно использовать при подборе значения интенсивности в точках D, E, даже если в конкретное время у нас не было получено значительного числа треков, проехавших через данный перекресток.

Первое предположение основывается на том, что в районе регулируемого перекрестка ТП становится более детерминирован. Имея модель движения ТП, при

известных схемах организации движения на перекрестках(схемы пофазного разъезда, длительности и смещение фаз), фиксированных значений пропускных способностей, вероятно можно тем или иным способом подобрать такие значения интенсивности в узлах графа, чтобы модельные треки движения ТС между точками D и E были “максимально близки” к GPS-трекам.

Если рассматривать каждый трек, ограниченный определенным участком дорожной сети, как точку в некотором пространстве, то можно попытаться разработать подходящую метрику, определяющую расстояния между двумя точками этого пространства. Это позволит сравнивать проекцию GPS-треков на их маршрут по графу дорожной сети (long, lat, time) в (time, distance) и треки, построенные моделью. Далее под измеренным GPS-треком понимается именно такой, спроецированный на граф участка дорожной сети в координатах (time, distance).

Второе предположение основывается на наблюдении о том, что часто, к примеру утром, многие едут по направлению к центру города, а вечером на том же перекрестке уже более нагружено обратное направление. И текущий 10 минутный временной интервал можно будет отнести к той или иной известной группе временных отрезков, где будет достаточное количество треков и рассчитывать на то, что закон больших чисел не подведет.

3. Модель транспортных потоков

Модель представляет собой направленный граф, вершины которого обозначают некоторые(важные) точки на дорожной сети, дуги это связи существующие между этими вершинами, определяемые структурой дорожной сети. Есть несколько типов вершин, каждая вершина обладает определенным набором свойств, например: интенсивность, пропускная способность дороги, идентификатор, время проезда и тд.

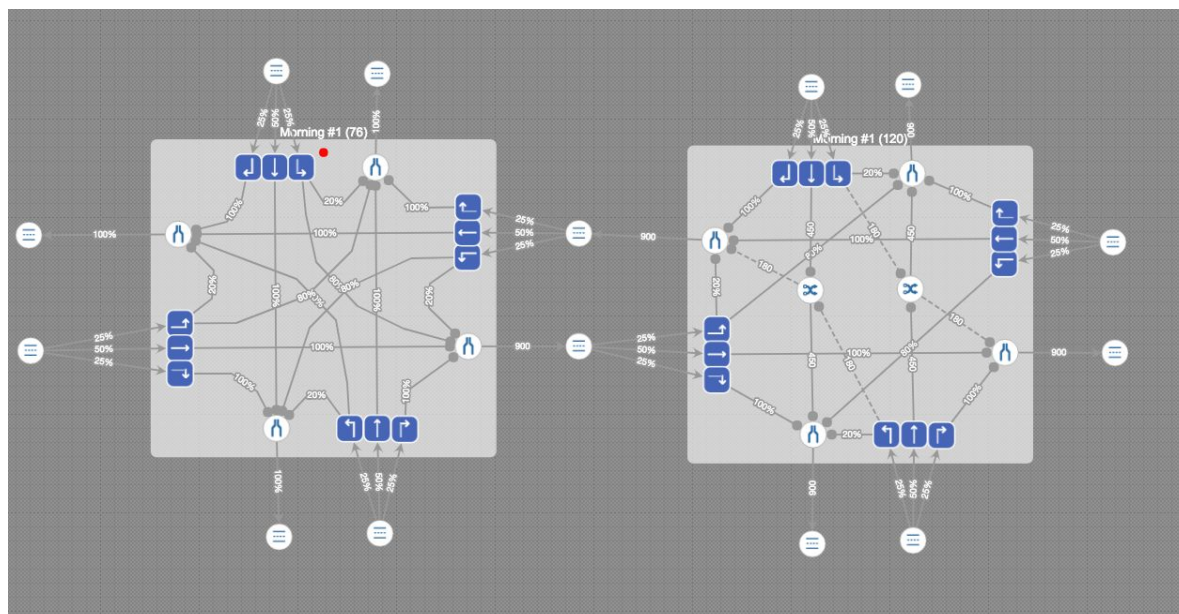


Рис.2 Пример модели описывающий 2 смежных перекрестка, “стоп линия” - синий знак, “бут. горлышко” - воронка, “перегон” - значок в виде дороги.

Вершина является функциональным элементом модели, с одним или несколькими входами и одним или несколькими выходами. На вход подается функция интенсивности от времени (временной профиль потока) в пределах длительности цикла моделирования.

Каждая вершина в зависимости от своего типа, тем или иным образом преобразует профиль потока на ее входе(ах) и возвращает новый на своем выходе. Существуют следующие типы вершин:

- автомобильная стоп линия;
- перегон;
- конфликтная точка;
- слияние с приоритетом;
- бутылочное горлышко;
- пешеходная стоп линия

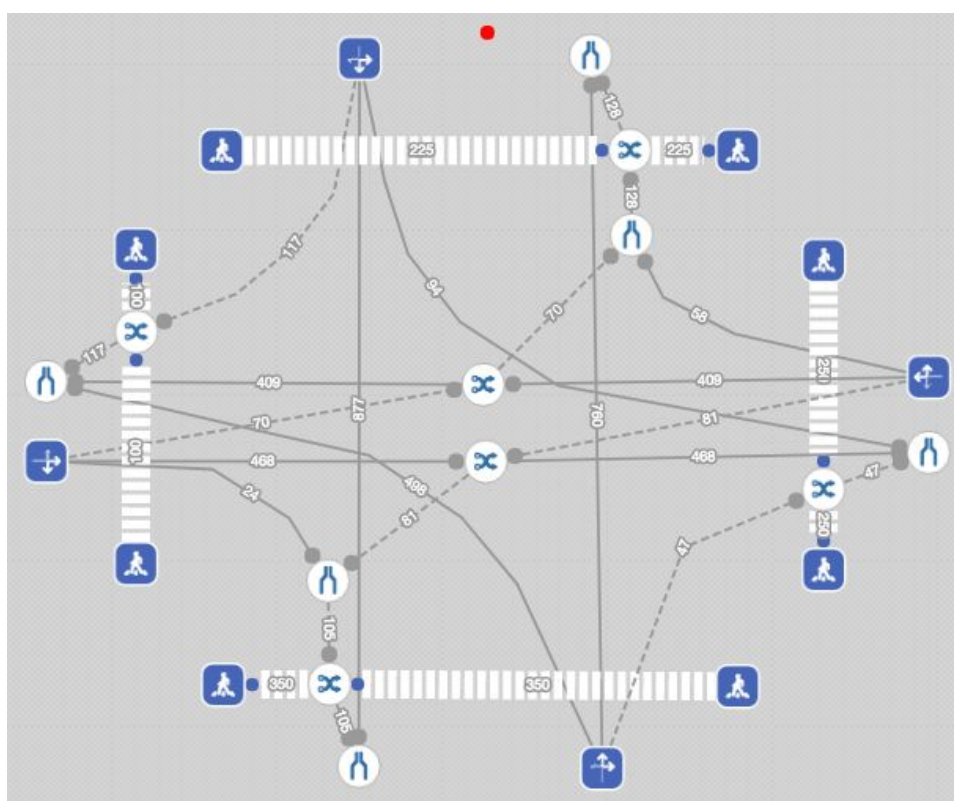


Рис. 3 Граф описывающий движение и взаимодействие ТП на перекрестке, числа на дугах обозначают интенсивность потока в авт/ч. Пунктирная дуга - второстепенный (просачивающийся) поток.

Стоп линия.

На вход требуются такие параметры как, приведенная к легковым интенсивность Q (авт/ч), пропускная способность C (авт/ч), длительность цикла моделирования (сек), длительность цикла светофорного регулирования T (сек), массив временных интервалов горения красного сигнала $[[0,10],[30,45]]$ и функция(ии) изменения интенсивности(профиль потока) в течении цикла регулирования с предыдущих точек - $q(t)$ (авт/с), если нет таких то она устанавливается как $q(t) = \text{const}$. Интервал дискретизации $\Delta t = 1$ сек.

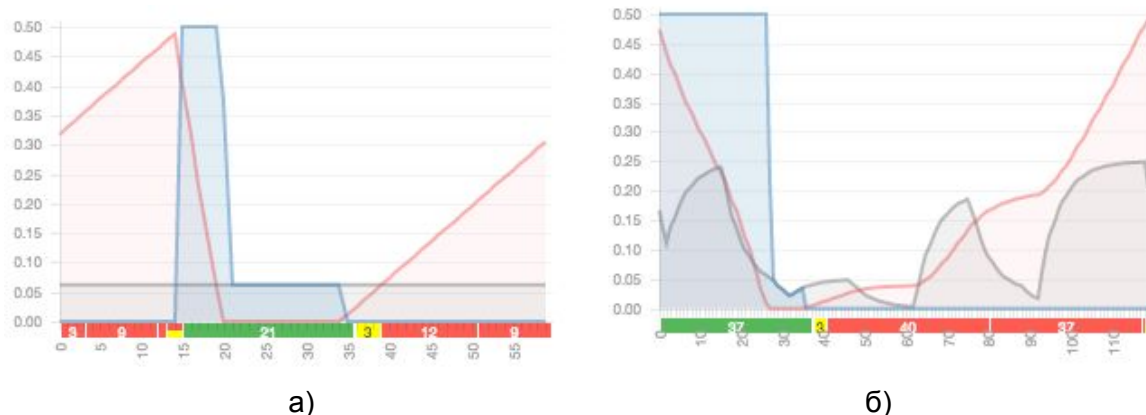


Рис 4. Профиль потока в течении цикла регулирования на стоп линии перекрестка $q(t)$. Серая линия входной профиль $q_i(t)$. Синяя - выходной $q_o(t)$. , красная - длина очереди в авт (на графике нормирована по пропускной способности). а) - $q(t) = \text{const}$ б) $q(t)$ - берется с предыдущей(их) смежной вершины. Ось x - время в секундах, ось y - интенсивность авт/с.

Перегон

На вход требуются такие параметры как, приведенная к легковым интенсивность Q (авт/ч), пропускная способность C (авт/ч), длительность цикла моделирования (сек), длительность цикла регулирования T (сек), Функция(ии) изменения интенсивности в течении цикла регулирования с предыдущих точек - $q(t)$. В отличии от стоп линии тут требуются два дополнительных параметра. Длина - L в метрах и время проезда T_r (сек).

В данной модели используется вертикальная очередь, то есть формула распада пачки[1] не учитывает, того что длина перегона изменится. Получается, что на следующей стоп линии очередь выстраивается не назад, изменяя длину перегона, а вверх.

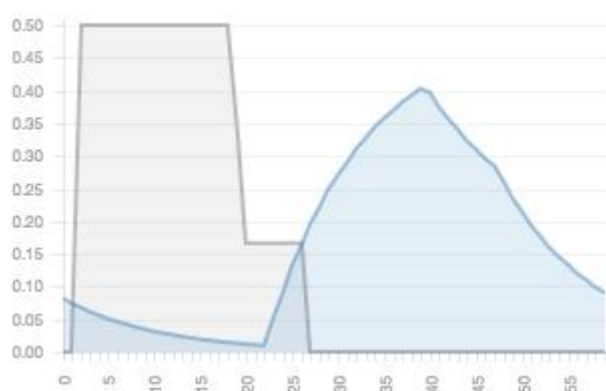


Рис. 5 Функция(ии) изменения интенсивности в точке типа перегон в течении цикла регулирования. Серая линия входной профиль потока $q_i(t)$. Синяя - выход $q_o(t)$. Ось x - время в секундах, ось y - интенсивность авт/с.

Если узлы типа стоп линия и бутылочное горлышко можно отразить на реальную дорогу в виде, некоторой линии, которую пересекают автомобили, то точка

перегон отражается в участок дороги который имеет уже длину, время проезда(скорость движения) они задаются в качестве входных параметров модели.

В связях между точками указывается значения как перераспределяется интенсивность из одного узла в другой. Подавая на вход модели все указанные параметры мы можем рассчитать функцию длины очереди на всех стоп линиях и других точках, где она может возникнуть. И рассчитать задержку в авт*с. На базе задержки и весов узлов рассчитывать и оптимизировать параметры регулирования на перекрестках, длительности фаз, относительные сдвиги при координированном управлении.

Бутылочное горлышко, выход из перекрестка.

Рассмотрим точку еще одного типа, представление о которой будет важно для понимания следующих частей работы. Точка типа “бутылочной горлышко” используется в модели для описания сужений проезжей части или места где несколько потоков могут сливаться в один и может возникнуть очередь, в частности на выходе из перекрестка.

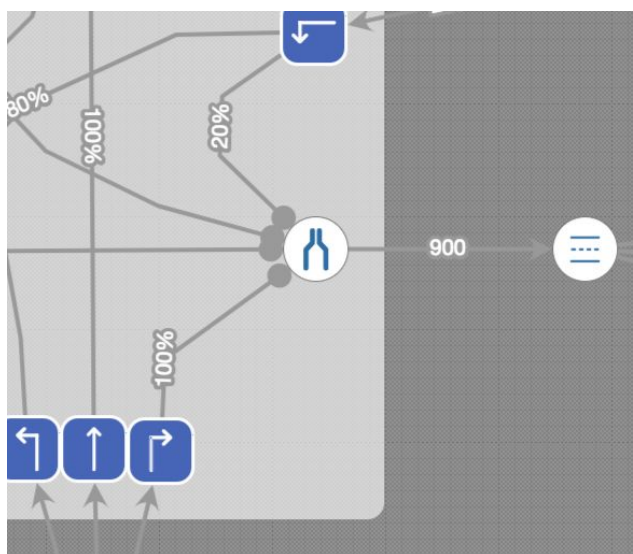


Рис.6 Точка типа бутылочное горлышко

Профили потока из соседних вершин посекундно складываются между собой, если будет превышено значение пропускной способности, то формируется очередь и рассчитываются задержка(см рис 7а). Важно понять, что в этой точке мы можем увидеть как складываются во времени потоки из разных фаз движения на перекрестке, например, на рисунке 7б показано, что на указанный выход из перекрестка движение разрешено только в 1 и 2 фазе, в остальное время значение потока равно нулю.

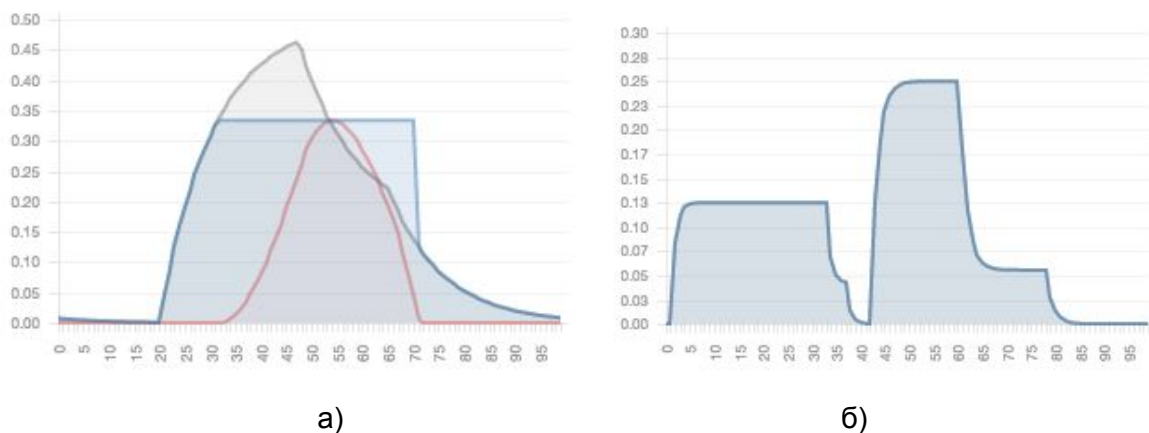


Рис. 7 Профили потока в точке бутылочное горлышко. а) Сужение с образованием очереди. б) в точке выхода 3х фазного перекрестка, слияние 2х потоков, разделенных во времени светофорным регулированием. Ось x - время в секундах, ось y - интенсивность авт/с.

Модельные треки передвижения автомобилей

Складывая значения профиля потока, в точке выхода из перекрестка, можно получить время, когда каждое ТС покинет перекресток. Так как профиль потока в модели имеет единицы измерения авт/с, то последовательно складывая его значения и отмечая, те моменты времени, когда значение суммы будет переходить через целые числа, можно получить время выхода каждого ТС. Начиная с этих времен выхода строим треки.

Модель построения трека самая простая: можно ехать - едет, нет - стоит. Все едут равномерно, ускоряются мгновенно, очередь горизонтальная, габарит ТС в очереди 5м.

Модельные треки, а также времена выхода из перекрестка, могут быть получены с помощью *любой другой* подходящей модели движения ТП, и скорее всего треки полученные микро-моделированием, будут гораздо точнее отражать реальное перемещение ТС. Такую упрощенную модель, как описано выше, проще запрограммировать и быстрее рассчитывать. В такой ситуации просто необходимо выбрать из соотношения возможность-скорость-точность.

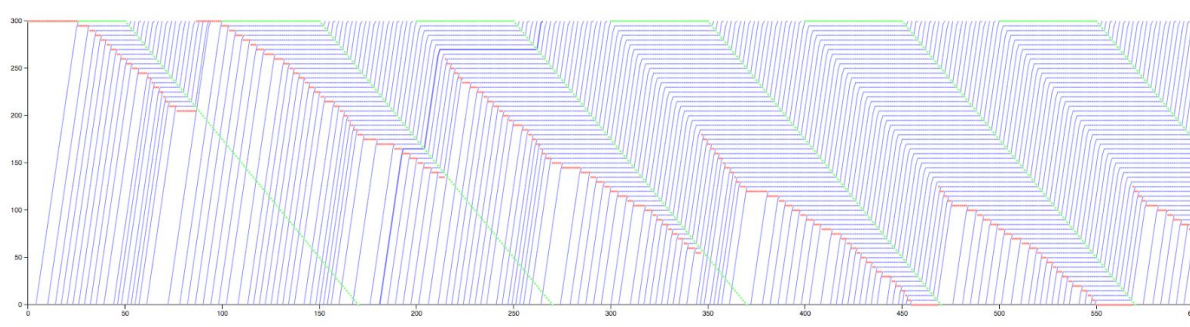


Рис 8. Модельные треки движения по перегону с учетом профиля потока на входе и светофорного регулирования на следующей стоп линии. Ось x - время в секундах, y - расстояние в метрах.

4. Исходные данные

В численных экспериментах описанных ниже в качестве измеренных треков выступают треки построенные в координатах $time$, $distance$ с помощью описанной выше модели. Шаг дискретизации по времени равен 1 секунде, по расстоянию 1 метр. Как было сказано выше, ускорение ТС в модели построения трека стремится к бесконечности, поэтому на графике (на рис 8) треки выглядят, как ломаная линия состоящая из отрезков движения и отрезков, когда координата $dist$ не изменяется, те ТС стоит.

Понятно что, с реальными GPS - треками, все будет обстоять несколько иначе и описанные выше характеристики трудно достижимый идеал, как по дискретизации так и по точности.

Для подтверждения [предположения №2](#) не обязательно иметь большую точность позиционирования во времени и пространстве. Для всех измеренных треков, важно их количество и достоверно знать только через какие входы выходы перекрестка проехало ТС. Для [предположения №1](#), напротив, точность измерения точек трека будет иметь очень важное значение.

В настоящее время очень много автомобилей спецтехники, общественного транспорта и маршрутных такси оборудованы трекерами, характеристики которых значительно лучше чем обычных мобильных телефонов. Совместное использование треков от ТС разного типа позволит улучшить точность оценки.

При данном подходе не рассматриваются треки проехавшие без остановки, однако по ним можно судить о таком параметре модели, как скорость/время свободного проезда.

Очевидно, что модель построения треков достаточно простая, в реальности линии треков на графике будут выглядеть более плавно (ускорение, $v \neq const$), однако, возможно не стоит "тянуть" модельные треки к реальности (повышать адекватность модели), вероятно дешевле в вычислительном плане будет аппроксимировать точки реальных треков к виду упрощенных модельных. Тем более, как это будет показано ниже, некоторые признаки треков опираются на точки переломов, определять которые по гладким кривым будет несколько сложнее.

5. Метрика, пространство признаков треков

Первое, что приходит в голову и очевидно самой лучшей мерой близости 2-х треков, будет площадь криволинейной фигуры образованной этим 2-мя треками и 2-мя горизонтальными линиями $y=0$ и $y=d$, где d - расстояние от выхода первого перекрестка до стоп линии второго перекрестка (d - длина дуги DE см рисунки 1, 9а). Такая мера близости хорошо подходит для сравнения треков между собой и ответа на вопрос насколько один трек отличается от другого, но не всегда достаточна для определения интенсивности. На рисунке 9б пример из двух пар треков, площади фигур которых равны $S_1=S_2$, но невозможно достоверно сказать, что интенсивности в случае красных треков, отличается пропорционально значению S от черных в обоих случаях.

Поэтому было принято решение собрать максимально возможный набор признаков, которыми может характеризоваться данный трек и попробовать строить метрику на базе этого набора.

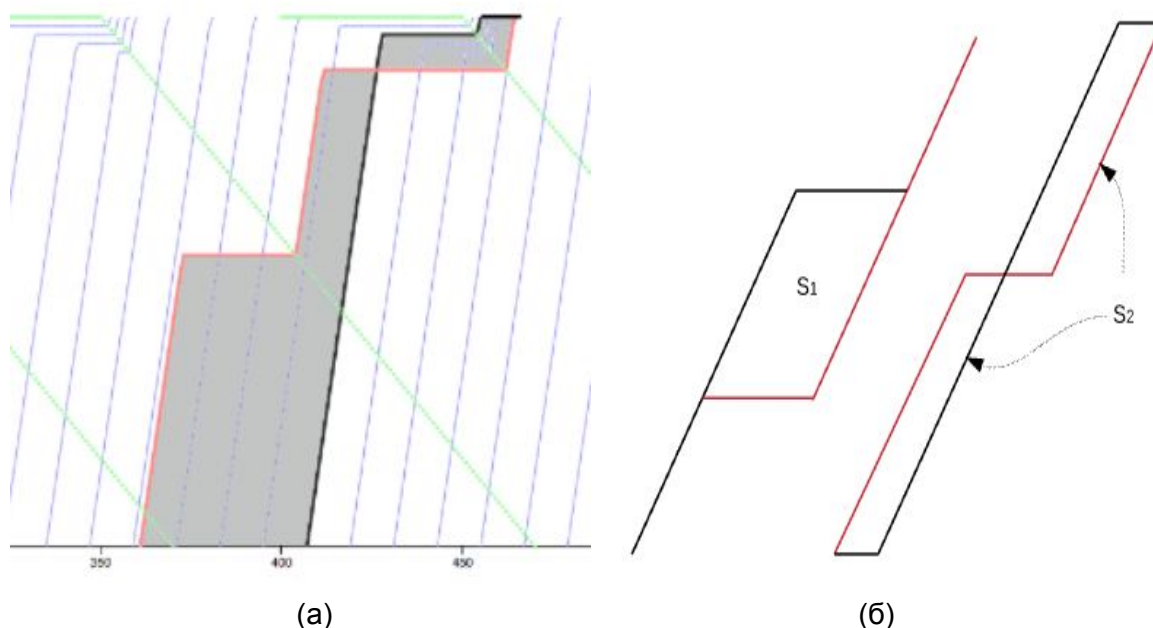


Рис. 9 (а) Графическое представление меры близости 2-х треков. Ось x - время в секундах, y - расстояние в метрах. (б) треки с одинаковой площадью $S_1 = S_2$, но отражающие различную дорожную ситуацию.

Были выбраны следующие признаки:

1. T_R - время проезда трека по дуге DE в секундах;
2. T_S - время проведенное в ожидании, скорость = 0;
3. S_L - площадь прямоугольника(ов), где одной стороной является время ожидания в очереди, а второй расстояние до стоп линии во время этого ожидания;
4. $\tan(a)$ - тангенс (или значение в градусах) угла образованного между двумя прямыми. Первая это прямая проведенная между двумя точками, (t_R, d) и точкой, остановки ТС, вторая это прямая между точками (t_R, d) и (t_G, d) , где t_R - время включения запрещающего сигнала светофора; t_G - время включения следующего разрешающего сигнала светофора; Если ТС ожидало возможности проезда несколько раз, то сумма углов складывается.

Значение угла a ($\tan(a)$) является важным признаком, так как если бы поток был равномерным, и точно определить точку остановки ТС, то вполне хватило бы его одного, чтобы рассчитать интенсивность, но в реальности, это конечно маловероятно. На рисунке 10 описываемые признаки представлены графически.

В качестве метрики в численных экспериментах использовалась евклидово расстояние между соответствующими признаками двух треков, но также в сумму под корень было введено значение площади между треками.

$$M = \sqrt{S(t_1, t_2) + (T_{R1} - T_{R2})^2 + (T_{S1} - T_{S2})^2 + (\sum S_{L1} - \sum S_{L2})^2 + (\sum \tan(a_1) - \sum \tan(a_2))^2}$$

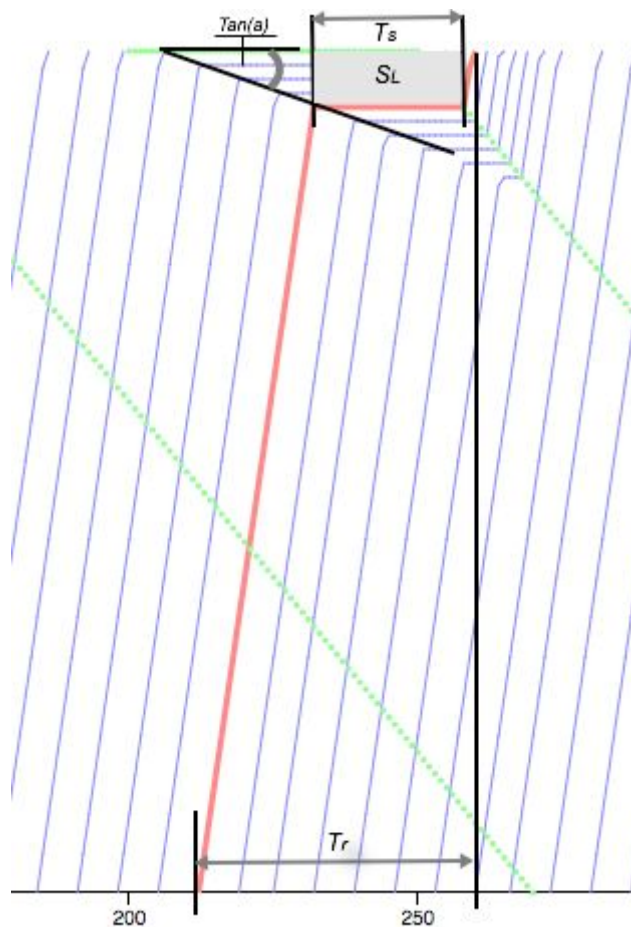


Рис. 10 Графическое представление признаков трека.

6. Код

Весь код необходимый для проведения экспериментов находится в репозитории на [github](https://github.com/kuzinmv/avenue). в папке [/backend/exp](#). Для того, чтобы попробовать самостоятельно не обязательно разворачивать весь проект, достаточно просто клонировать репозиторий или [скачать архив](#) и перейти в папку.

```
$ git clone https://github.com/kuzinmv/avenue.git
$ cd avenue
```

```
$ npm install
$ cd backend/exp
```

graph.js

Команда для построения и визуализации треков `graph.js` подготавливает данные для отрисовки графиков изображенных на рисунках 8, 9а,10. Параметры команды `graph.js`

```
$ node graph.js {file-name} {highlight-track-num} {overwrite-flow}
> out/traces-data.out.js
```

, где `{file-name}` - имя файла модели, `{highlight-track-num}` - номер трека, который надо выделить красным цветом(необязательный), `{overwrite-flow}` - переопределить значение интенсивности в модели и использовать равномерный поток(необязательный). Примеры вызова команды:

```
$ node graph.js data/exp-600.json 20 > out/traces-data.out.js
$ node graph.js data/exp-600.json 20 400 > out/traces-data.out.js
```

Далее достаточно просто открыть/обновить в браузере файл `out/view-traces.html` для просмотра получившихся треков в графическом виде.

single.js

Команда для построения таблицы признаков треков. В данной команде имеется возможность выбрать один необходимый трек по результатам моделирования и сравнить его с ближайшими по площади треками для всех значений интенсивности из диапазона [200, 1000] с шагом 2 авт/ч; В результате исполнения команды формируется csv файл. Параметры команды `single.js`

```
$ node single.js {file-name} {overwrite-flow} {first-tt}
```

, где `{file-name}` - имя файла модели, `{overwrite-flow}` - переопределить значение интенсивности в модели, `{first-tt}` - время в секундах, после которого будет взят первый подходящий трек, для сравнения со всеми остальными картинками треков при переборе значений интенсивности. Примеры вызова команды.

```
$ node single data/exp-full.json 500 200 > out/single-data.csv
```

На рисунке 11 приведены в виде графиков некоторые признаки треков в результате работы команды `single`.

compre.js

Команда для проведения полного численного эксперимента. Для каждого значения из набора интенсивностей проводится моделирование, и построение треков. Далее выбираются все подходящие под условия предположения 1 треки. И по каждому из них подбирается такое значение интенсивности, где значение *M* будет минимальным. Примеры вызова команды.

```
$ node compre data/exp-full.json > out/comprehensive-data.csv
```

В результате формируется таблица из двух колонок, в первой исходное значение интенсивности, во второй - значение определенное/подбранное по треку, исходя из минимума евклидового расстояния между признаками.

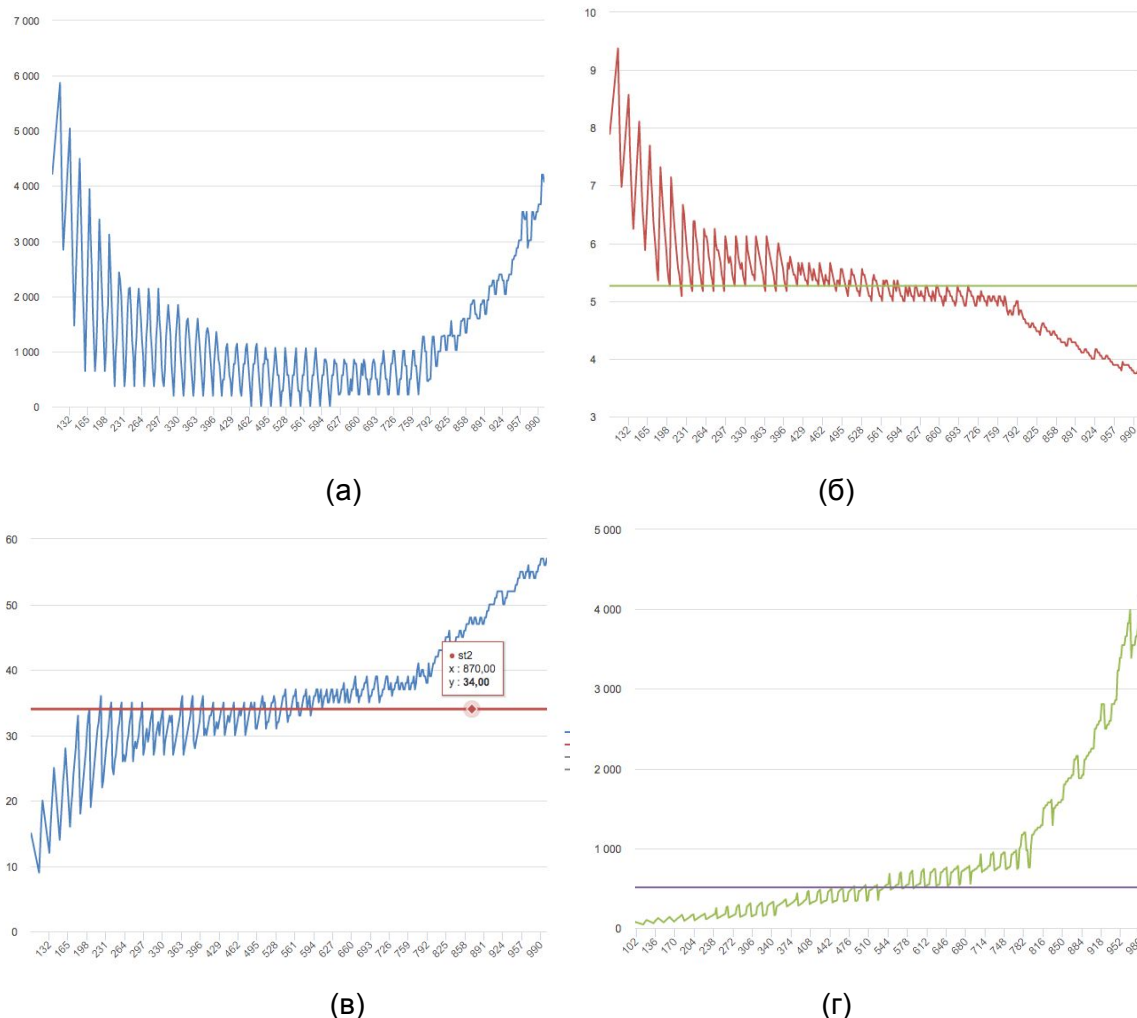


Рис 11. Результаты работы команды single. Зависимости признаков выбранного трека и ближайшего к нему от интенсивности(ось x). а) S - площадь м/у треками; б) v - скорость проезда (прямая линия - выбранный трек, кривая. - ближайший к нему); в) время простоя в очереди; г) площадь прямоугольника(ов) S_L .

7. Результаты экспериментов с моделью

В результате экспериментов с моделью, практически сразу выяснилось, что [предположение 1](#) неверно, и одного трека будет недостаточно. На рисунке 12 приведен график значения площади S между выбранным треком и ближайшим при заданной интенсивности (ось x на графике). В районе значения 500 (настоящее значение) в диапазоне примерно (450, 620) находится 9 точек, значение которых равно 0. Это означает то, что трек не уникальный и находятся треки точка в точку совпадающие с выбранным. Причем правильное значение интенсивности часто

находится ближе к краю этого интервала и любое усреднение дает значительную ошибку.

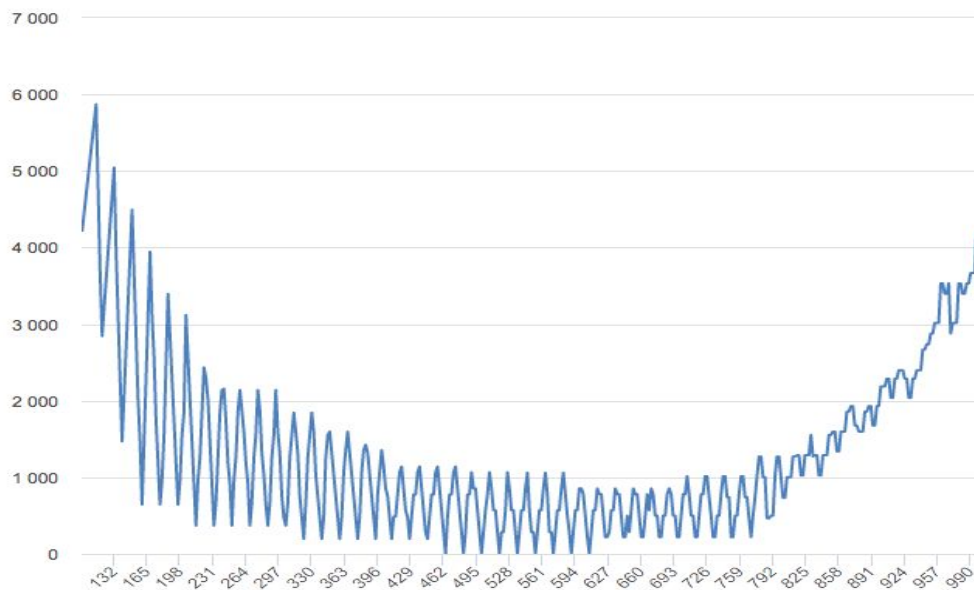


Рис. 12. Значения площади S между выбранным треком и ближайшим при заданной интенсивности (ось x на графике)

На рисунке 13 приведены примеры идентичных треков полученных при значении интенсивностей 220 и 440 соответственно. Такие треки - это чаще всего первые или вторые автомобили в очереди, наиболее широкий диапазон ошибки получается в средних значениях интенсивности 300-700, в верхних и нижних диапазонах такое наблюдается гораздо реже.

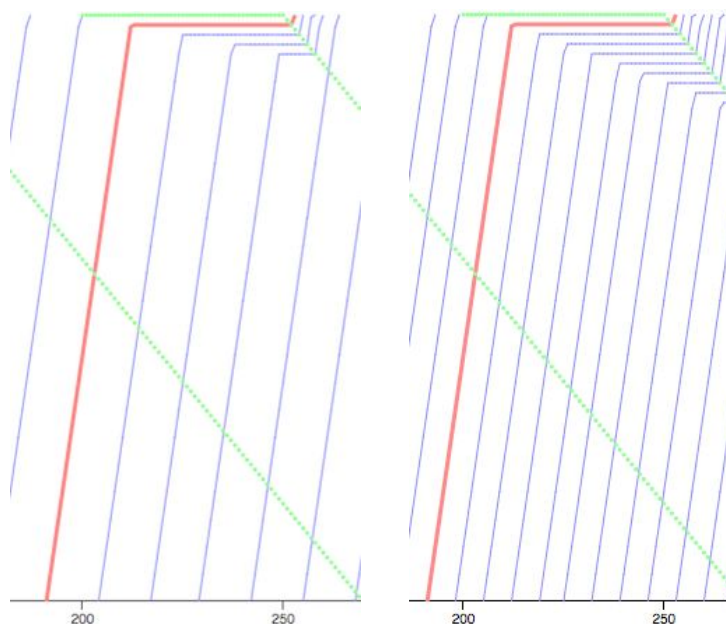


Рис. 13. Абсолютно идентичные треки при разных значениях интенсивности

Перестроим полный численный эксперимент следующим образом. Зафиксируем значение интенсивности Q_1 и построим картину треков, выберем из них только подходящие. Для каждого трека из этого множества, случайным образом выберем пару, таким образом у нас будет два трека. Найдем все точки $(Q, M = \min(M))$, которые имеют минимальные значения метрики M для обоих треков, и возьмем среднее значение координат Q этих точек по пересечению значений Q из получившихся массивов. Обозначим получившееся в результате значение как Q_2 - интенсивность определенная по двум трекам.

На рисунке 14 результаты 2-х экспериментов, значение Q_1 (ось x) задавалось в диапазоне $[200, 900]$ с шагом 100(а) и 50(б), по оси y отмечаем значения Q_2 .

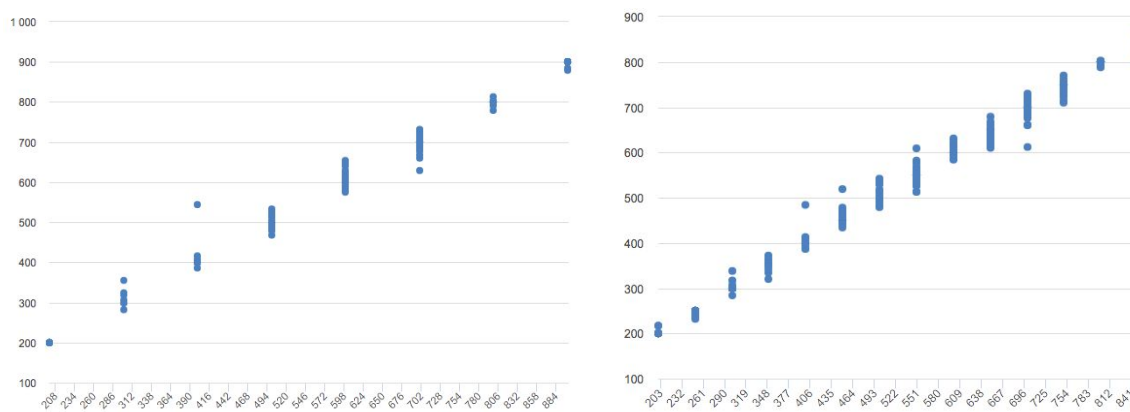


Рис 14. Результаты полного эксперимента по 2м трекам.
а) шаг $dQ_1 = 100$, б) шаг $dQ_1 = 50$

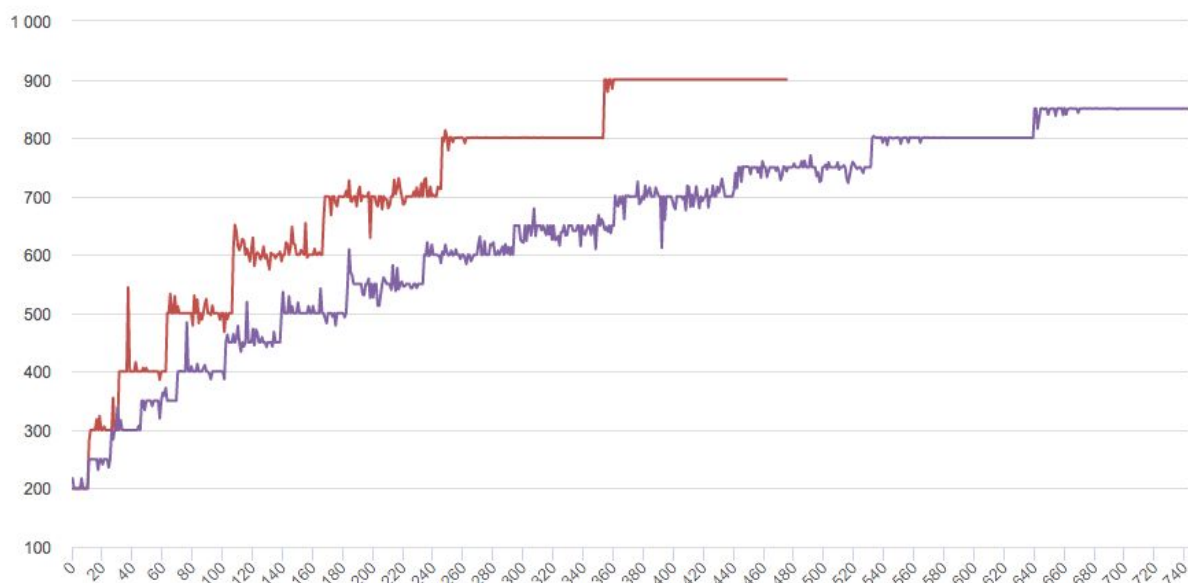


Рис. 15 Те же два эксперимента, но по оси x откладываем номер трека(чем больше Q_1 тем шире полочка - больше подходящих треков), по оси y значение Q_2 определенное для данного трека и его случайной пары.

8. Заключение.

Исходя из результатов экспериментов, виртуальный детектор транспорта кажется вполне возможным. Если даже такая, практически, первая попавшаяся метрика дает неплохие результаты всего по двум трекам. При том, что даже не используются никакие признаки, относящиеся к группе треков, а только к одному. Метрика учитывающая сразу несколько треков а не только пару, должна уменьшить ошибку.

В данной статье рассматривается пример, где интенсивность определяется только на одном выходе из перекрестка. Образно выражаясь идет попытка подобрать корень только одного уравнения, однако все выходы из перекрестка связаны распределением вероятностей разъезда/съезда на входах/выходах и решая их одновременно, как "систему уравнений" перекрестка, могут быть получены более точные результаты.

Проверка предположения 2 не производилась и она возможна только по реальным данным, возможна ли такая кластеризация, насколько будут четкие границы кластеров, какие признаки использовать - все это еще требует тщательного исследования.

В данном исследовании рассматривался пример, и модель построения треков, только на однополосной дороге, многополосность в любом случае внесет свои коррективы, как в признаки, так и в метрику;

Так же пока за рамками оставлен момент с "периодичностью" значения многих признаков треков и возможным влиянием значений/изменения этого периода на алгоритм.

В рассмотрении могут участвовать не 2 смежных перекрестка, а 3 и более.

Значения в интенсивности полученные, на одном, самом "красивом" перекрестке, со множеством треков во всех направлениях в заданное время, могут быть использованы в подборе на соседних перекрестках улучшая точность и скорость.

Интенсивность предполагается получать только путем подбора, на основе полученных за некоторое короткое время измерений треков проезда, но нам ничего не мешает предварительно ввести в модель известные данные об интенсивности снятые со стационарных детекторов или же наколенные описанным выше способом. История интенсивностей пусть даже с ошибками, начнет влиять на модель оценки и сделает ее точнее.