

## **ИСПОЛЬЗОВАНИЕ КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ И МАТЕМАТИЧЕСКИХ МЕТОДОВ ДЛЯ АНАЛИЗА АВТОТРАНСПОРТНЫХ СЕТЕЙ**

**А.М. Пуртов**

доцент, к.т.н., с.н.с. ОФ ИМ СО РАН, e-mail: andr.purtov@yandex.ru

**Аннотация.** Разработан способ применения таксономии, редукции графов и методов геоинформационных систем для анализа влияния задержек на время прохождения маршрутов в автотранспортных сетях. Технология демонстрируется на примере анализа популярного маршрута г. Омска. Приведён пример использования таксономии для анализа результатов редукции графа. Показано сходство результатов визуальной и автоматической таксономии. Результаты таксономии отображены на ГИС – карте графа маршрута. Разработана имитационная модель перекрёстка. Модель запрограммирована на GPSSW.

**Ключевые слова:** геоинформационная система, имитационное моделирование, редукция графов, таксономия, анализ данных, автотранспортные сети.

*Работа выполнена при поддержке гранта РФФИ 12-07-00149-а*

### **Введение**

Рост возможностей приобретения автомобилей населением приводит к перегрузке автотранспортных сетей. Поэтому задачи анализа и организации автотранспортного движения становятся все более актуальными. Современные компьютерные технологии и математические методы имеют большие возможности для решения такого рода задач. Интеграция различных методов и средств позволяет с разных сторон посмотреть на возникающие проблемы. Создаваемая автором система GisAuto ориентирована на анализ времени прохождения различных участков пути. В GisAuto предполагается совместное использование методов геоинформационных систем (ГИС), математических методов, в частности, методов редукции графов, таксономии, имитационного моделирования. В [1–3] более подробно, чем в предлагаемой статье, показана эффективность интеграции перечисленных методов. Начало широкого использования ГИС для отображения и решения проблем на транспорте можно отнести к концу прошлого века. Представляемая работа ориентирована на анализ времени прохождения маршрутов с целью оперативного управления транспортными потоками и стратегического планирования автотранспортной сети. Суть разрабатываемой технологии можно отобразить следующими этапами.

1. Построение ГИС-модели задержек на основных маршрутах города. Задержки происходят на перекрёстках, светофорах, пешеходных переходах.
2. Построение на ГИС-карте графов исследуемых маршрутов.
3. Сбор данных о задержках. На этом этапе могут быть использованы экспертные, расчётные оценки, результаты наблюдений, имитационного моделирования.
4. Анализ маршрутов методом редукции графов. В результате получают коэффициенты, показывающие влияние каждой задержки на общую задержку при прохождении маршрута.
5. Раскраска задержек (вершин графа) на ГИС-карте в зависимости от их значимости.
6. Анализ полученных результатов. Выявление задержек и участков, оказывающих наибольшее влияние на время прохождения маршрута.
7. Микроанализ проблемных участков (дополнительные наблюдения, имитационное моделирование).

В статье на примере анализа одного из популярных маршрутов г. Омска показано выполнение перечисленных этапов. Новизна разрабатываемой технологии заключается в оригинальном сочетании разработанных автором моделей, реализованных различными методами.

## **1. Построение ГИС-модели задержек и графов маршрутов**

ГИС-модель задержек на дорогах разрабатывается с целью предоставления первичных данных для анализа транспортных сетей различными методами (аналитическое и имитационное моделирование, геоанализ). Подробно ГИС-модель задержек и построение графов маршрутов описаны в [1]. Основными объектами модели являются задержки и участки дорог. Выделены следующие типы задержек: перекрёсток, светофор, пешеходный переход. Остановки пассажирского транспорта в модели не учитываются. Если на переходе имеется светофор, задержка перед ним относится к типу «светофор». Если перекрёсток оборудован светофорами, задержки на нем относятся к типу «перекрёсток». Участки дорог соединяют задержки. На основе ГИС-модели задержек строятся графы маршрутов. Для перемещения между двумя точками могут использоваться разные пути. На основе множества основных путей получается граф маршрута. В левой части рис. 1 показаны основные задержки на дорогах г. Омска. В правой части рис. 1 показан граф одного из популярных маршрутов г. Омска (п. Солнечный — завод Баранова).

ГИС-модель задержек и граф маршрута реализованы средствами пакета GIS ArcView 3.0.

## **2. Анализ маршрутов методами ГИС и редукции графов**

Метод редукции графов для анализа вычислительных процессов в ЭВМ был предложен Б. Байцером [4]. Редукция заключается в свёртке (последовательном упрощении) графа с помощью замен фрагментов графа разработанными



Рис. 1. ГИС-карта задержек и граф маршрута

упрощающими подстановками. В дальнейшем метод Б. Байцера был модифицирован В.Н. Задорожным [5]. Примеры двух из восьми подстановок приведены на рис. 2.

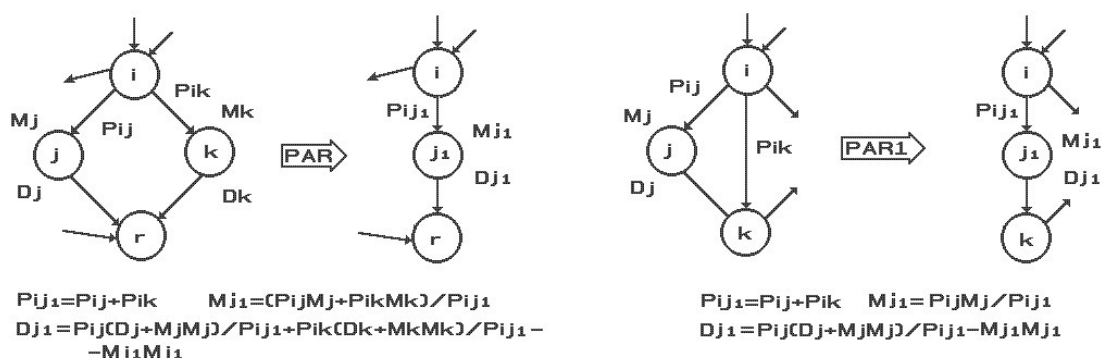


Рис. 2. Примеры упрощения и свертывания

Программная реализация метода была сделана автором статьи.

Суть метода заключается в следующем. Оценки математического ожидания времени ( $T_i$ ) и дисперсии ( $D_i$ ) задержек, вероятностей выбора в узлах путей следования ( $P_{ij}$ ) задаются на входе программы редукции графов. На выходе программы получают значения средней задержки при прохождении маршрута ( $T$ ), дисперсии  $T(D)$ , коэффициентов чувствительности  $T$  и  $D$  к параметрам задержек и вероятностям переходов (десять коэффициентов). Для получения оценки среднего времени прохождения маршрута необходимо к  $T$  добавить время свободного пробега (без задержек). Наиболее простую и физически объяснимую интерпретацию имеют абсолютный коэффициент чувствительности  $T$  к  $T_i$  ( $K_a[T, T_i]$ ) и относительный коэффициент чувствительности  $T$  к  $T_i$  ( $K_b[T, T_i]$ ). Дело в том, что при предположении о независимости параметров

исходного графа,  $K_a[T, T_i]$  представляет собой частную производную  $T$  по  $T_i$ , интерпретируемую как вероятность реализации задержки  $T_i$  при прохождении маршрута. Выполняется равенство:

$$T = \sum_{i=1}^n K_a[T, T_i]T_i, \text{ где } n \text{ — число вершин графа.}$$

Разделив обе части равенства на  $T$ , получаем сумму  $K_b[T, T_i]$ , равную 1. Таким образом, коэффициент  $K_b[T, T_i]$  показывает в долях от 1 вклад задержки  $T_i$  в общую задержку  $T$ . Учитывая свойства  $K_b[T, T_i]$ , будем называть его коэффициентом значимости задержки. Другие коэффициенты чувствительности не имеют такой простой интерпретации. Тем не менее, они иногда оказываются полезными для оценки влияния параметров  $T_i$ ,  $D_i$ ,  $P_{ij}$  на  $T$  и  $D$ . Ранее метод использовался автором статьи для анализа процессов передачи данных в компьютерных сетях [6]. Подробно применение метода для анализа автотранспортных сетей описано в [1].

Целью анализа является сравнительная оценка величин задержек и вероятностей выбора участков дорог с точки зрения степени их влияния на общую задержку при прохождении маршрута. Для достижения цели выполняются следующие действия:

1. Выбирается точка А (первая задержка на маршруте) и точка Б (последняя задержка на маршруте).
2. На ГИС-карте задержек выбираются те задержки и участки дорог, по которым обычно движется автотранспорт из точки А в точку Б (маршрут АБ).
3. Средствами GIS ArcView автоматически создаются слой задержек и слой участков дорог, относящихся к маршруту АБ. Слои отображаются на ГИС-карте. В результате получается граф маршрута АБ (рис.1).
4. Определяются (различными доступными методами) и записываются в базу данных оценки исходных параметров графа  $T_i$ ,  $D_i$ ,  $P_{ij}$ .
5. Исходные параметры графа подаются на вход программы редукции графов.
6. На выходе программы редукции графов получают оценки параметров  $T$ ,  $D$  и значения коэффициентов чувствительности.
7. Выбираются интересные для исследователя коэффициенты чувствительности. Создаются слои, на которых задержки и участки дорог изображаются в зависимости от значений соответствующих коэффициентов чувствительности.
8. Проводится геоанализ по всем полученным результатам (по результатам одного или нескольких экспериментов, по результатам анализа одного или нескольких маршрутов).
9. Результаты анализа представляются для обсуждения и принятия решений.

На рис. 3 показаны результаты анализа графа, представленного на рис. 1. Задержки изображены в зависимости от значения коэффициента  $K_b[T, T_i]$ .



Рис. 3. Коэффициенты значимости задержек на маршруте

В левой части рисунка показан весь маршрут, в правой части изображена центральная часть г. Омска. Как выше было сказано,  $K_b[T, T_i]$  в долях от единицы показывает вклад задержки в общее время прохождения маршрута. Поэтому  $K_b[T, T_i]$  называем коэффициентом значимости. На рис. 3 чем больше коэффициент, тем крупнее точка (данные разбиты на 3 категории). Кресты показывают задержки, оказывающие наибольшее влияние на время прохождения маршрута. Рисунок на уровне города наглядно показывает степень влияния каждой из задержек на общую задержку при прохождении маршрута.

Существуют задержки, значимые для нескольких маршрутов. На рис. 4 отображены коэффициенты  $K_b[T, T_i]$  участков маршрутов 1 и 2 в центре г. Омска. Маршрут 1 — от Радиостанции (на пр. Мира) до ж.д. вокзала. Маршрут 2 — от п. Солнечный до завода Баранова. Для маршрута 1 коэффициенты  $K_b[T, T_i]$  изображены белыми окружностями. Для маршрута 2 — крестами. Чем больше  $K_b[T, T_i]$ , тем крупнее символ. Из рис. 4 видно, например, что задержка на пересечении ул. Красный путь и ул. Фрунзе наиболее значима для обоих маршрутов.

### 3. Использование таксономии для анализа задержек

По классификации задач анализа данных, предложенной Н. Г. Загоруйко [7], таксономия (классификация, кластеризация, категоризация) заключается в разделении «объектов по схожести их свойств». При этом в шкале наименований каждая группа похожих объектов как-то обозначается. Основным научно-практическим направлением в таксономии является автоматизация классификации в многомерном пространстве параметров (признаков). В настоящее время нет универсального алгоритма таксономии, «который мог бы составить реальную конкуренцию человеческой способности к обобщению» [8].

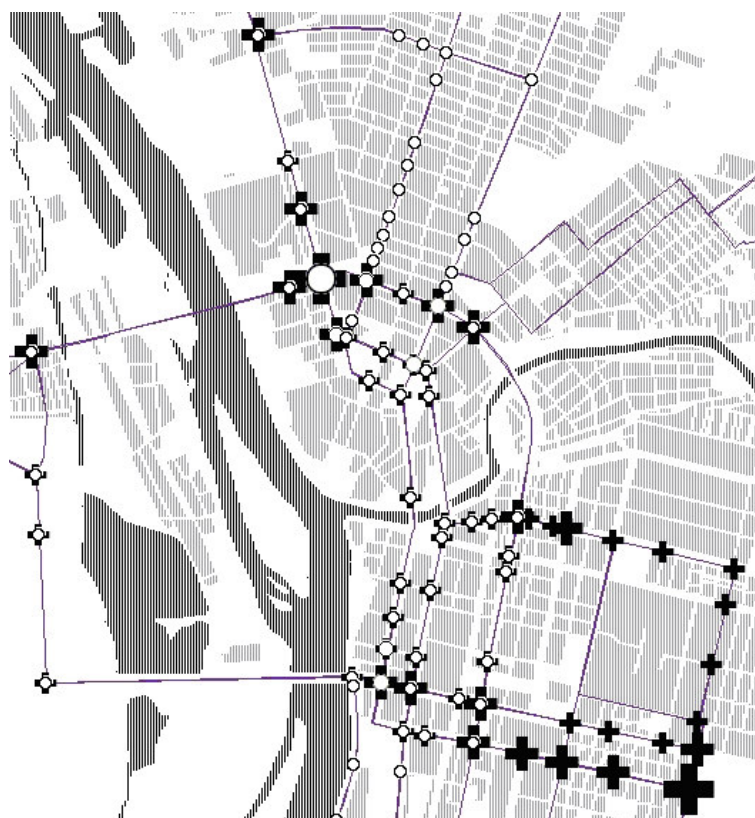


Рис. 4. Коэффициенты значимости задержек на двух маршрутах

Поэтому для таксономии в GisAuto используются автоматическая таксономия и предложенный автором статьи метод визуальной таксономии средствами ГИС [9].

Для автоматической таксономии студенткой СибАДИ Охотниковой К.В. при выполнении дипломной работы создана в среде C++ Builder программа, реализующая алгоритмы FOREL, FOREL-2, KOLAPS [7]. В статье приведены результаты использования визуальной таксономии и алгоритма FOREL-2, суть которого заключается в следующем.

В алгоритме FOREL-2 таксоны имеют форму гиперсферы (два параметра — круг, три параметра — шар). Параметры объектов считаются их координатами в  $n$ -мерном пространстве. Параметры нормируются, приводятся в единый диапазон, например, от 0 до 1. Считается, что чем больше объекты похожи друг на друга, тем ближе расположены в пространстве соответствующие им точки. Тогда задача таксономии состоит в том, чтобы выявить сгустки точек и объединить их в подмножества с похожими параметрами. В алгоритме FOREL-2 на входе задаются ограничение на допустимое количество таксонов, максимальный и минимальный радиусы гиперсферы. Оптимальным считается то разбиение, при котором минимальна сумма расстояний точек до центров таксонов.

При визуальной таксономии выбирается пара параметров объектов. Норми-

рованные параметры считаются координатами объектов на плоскости. Средствами ГИС с помощью геометрических фигур произвольной формы (круг, прямоугольник, многоугольник), используя выбранные критерии, точки, соответствующие объектам, объединяются в таксоны.

Основное преимущество автоматической таксономии по сравнению с визуальной заключается в возможности выявления похожести объектов в  $n$ -мерном пространстве. Кроме ограниченности двумерным представлением параметров к недостаткам визуальной таксономии можно отнести субъективизм разбиения на таксоны. С другой стороны, визуальная таксономия имеет следующие преимущества по сравнению с автоматической:

- возможность использования эвристических процедур при разбиении на таксоны;
- быстрый переход от одних критериев классификации к другим;
- более простая и логичная интерпретация результатов таксономии.

На рис. 5 представлено визуальное и автоматическое разбиение задержек на таксоны. Пунктирные линии обозначают вычисленные средние арифметические значения параметров задержек  $T_i$  и  $K_b[T, T_i]$ . Визуальное разбиение на таксоны показано геометрическими фигурами. При разбиении использовались критерии близости (кучности, похожести) и величины отклонения от среднего. Использование средних значений упрощает интерпретацию результатов. Например, в первый таксон попали задержки со средними значениями параметров, во втором таксоне — задержки с высокими значениями параметров, в третьем таксоне — задержки с высоким значением  $T_i$  и средним значением  $K_b[T, T_i]$ . Таксон 3 показывает, что не все большие задержки существенно влияют на время прохождения маршрута. Это из-за малой вероятности реализации этих задержек.

Результаты автоматической таксономии показаны разными формами точек. Например, мелкие точки попали в один таксон, треугольники в другой. В левой части рис. 5 представлены результаты автоматической таксономии по алгоритму FOREL-2 с использованием параметров  $T_i$  и  $K_b[T, T_i]$ . Возник вопрос о том, как изменятся результаты таксономии, если обработать алгоритмом FOREL-2 три параметра. На вход программы таксономии были поданы значения параметров  $T_i$ ,  $K_b[T, T_i]$  и  $K_a[T, T_i]$ . В правой части рис. 5 показаны результаты таксономии в плоскости параметров  $T_i$  и  $K_b[T, T_i]$ . В результате учёта параметра  $K_a[T, T_i]$  в число влиятельных добавились три задержки (обозначенных квадратами), имеющие высокую вероятность реализации.

На основе результатов, представленных на рис. 5, можно сделать выводы:

- в рассматриваемом случае визуальная таксономия мало отличается от автоматической;
- таксономия по двум параметрам мало отличается от таксономии по трём параметрам.

На рис. 6 показано изображение задержек на маршруте в зависимости от номера таксона, в который они попали при визуальной таксономии.

Кресты и крупные точки показывают наиболее значимые задержки (таксоны 2 и 5). Квадраты (таксон 3) иллюстрируют задержки с большим  $T_i$ , но редко

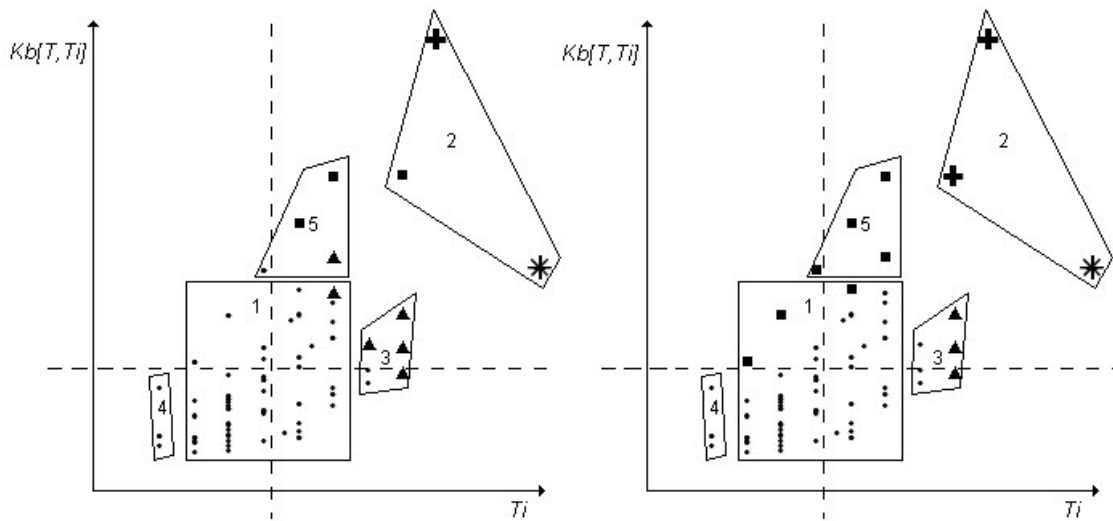


Рис. 5. Визуальная и автоматическая таксономии

реализуемые. Мелкие точки (таксоны 1 и 4) обозначают задержки, относительно слабо влияющие на прохождение маршрута. Разумеется, здесь речь не идёт о спонтанно возникающих на дорогах пробках. Изображение задержек и их значимости на реальной карте позволяет использовать эвристические методы анализа данных и принятия решений при развитии транспортной сети. Более подробно использование таксономии в GisAuto изложено в [2].

#### 4. Имитационная модель перекрёстка

Имитационное моделирование может использоваться как при оценке исходных данных для указанных выше методов, так и для детального анализа полученных результатов. Разработанная концептуальная модель перекрёстка демонстрируется на рис. 7. Реальные перекрёстки имеют множество разных конфигураций, полос и направлений движения, способов размещения светофоров. Предлагаемая модель достаточно универсальна. Она позволяет с помощью упрощений отобразить большинство перекрёстков (за исключением самых сложных).

На перекрёстке в каждом направлении существует очередь, управляемая светофорами. В моделях на уровне классических систем массового обслуживания тонкости управления светофорами обычно не учитываются. В предлагаемой модели очереди зависят не только от задержек и интенсивностей поступления транзактов, но и от СУС. Предлагаемая модель не учитывает такие детали, как перестроение автомобилей в другой ряд, поведение водителей, ДТП. Для построения моделей более детальных, чем предлагаемая, используются другие методы, например, метод клеточных автоматов. Для поставленных в работе целей, связанных с оценкой задержек автомобилей перед перекрёстком, слишком детальные модели не требуются.





Рис. 6. Изображение результатов таксономии на графе маршрута

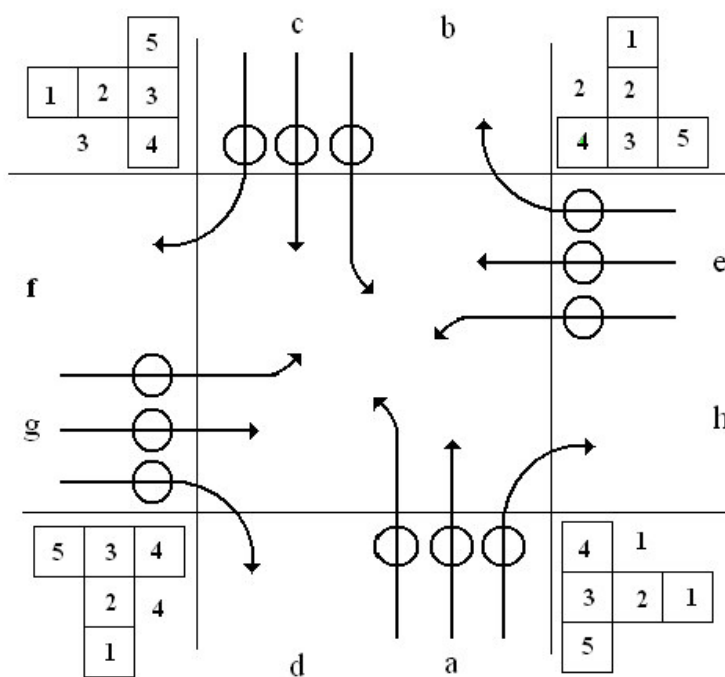


Рис. 7. Модель перекрестка

Интенсивность поступления автомобилей на перекрёсток (интенсивность генерации транзакций) в каждом из направлений зависит от предыдущей задержки (предыдущий светофор, переход). Концептуальная модель перекрёстка реализована автором статьи на GPSSW (студенческая, свободно распространяемая версия). Более подробно модель, а также проведённые эксперименты описаны в [3].

Разработанная имитационная модель может быть использоваться как в составе GisAuto для микроанализа перекрёстков, так и автономно, например, для сравнения алгоритмов управления светофорами и потоками автомобилей.

## Заключение

В статье на примере разрабатываемой системы GisAuto показана эффективность совместного использования ГИС-технологии и математических методов при анализе автотранспортных сетей. Дальнейшее развитие GisAuto связано с модернизацией разработанной технологии и включением в неё других методов.

## ЛИТЕРАТУРА

1. Пуртов А.М. Интеграция технологии ГИС и метода редукции графов для анализа транспортных сетей // Омский научный вестник. 2011. № 1(97). С. 164–168.
2. Пуртов А.М. Использование таксономии при анализе задержек в автотранспортных сетях // Вестник СибАДИ. 2013. № 2(30). С. 73–78.
3. Пуртов А.М. Разработка и анализ имитационной модели перекрёстка для системы GisAuto // Омский научный вестник. 2013. № 1(117). С. 225–229.
4. Байцер Б. Микроанализ производительности вычислительных систем: Пер. с англ. / Под ред. В.В. Мартынюка. М.: Радио и связь, 1983. 360 с.
5. Задорожный В.Н., Мызникова Т.А. Рекурсивный анализ чувствительности для метода Байцера. Деп. в ВИНТИ, 1988. № 5490-B88.
6. Пуртов А.М. Анализ производительности сетей ЭВМ на графах и имитационных моделях.: Автореф. дисс. на соиск. учен. степени канд. техн. наук (05.13.16) / Науч. рук. В.А. Шапцев. Омск: ИИТПМ СО РАН, 1995. 17 с.
7. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. Новосибирск: Изд-во Ин-та математики, 1999. 270 с.
8. Борисова И.А., Загоруйко Н.Г. Функции конкурентного сходства в задаче таксономии. Знания-Онтологии-Теории (ЗОНТ-07). Сб. науч. тр. Т. 2. Новосибирск, 2007. С. 67–76.
9. Пуртов А.М. Использование ГИС-технологии и таксономии для визуального анализа данных о субъектах РФ // Знания-Онтологии-Теории (ЗОНТ-09): материалы конф. с междунар. участием. Т. 2. (Новосибирск, 22-24 окт. 2009 г.). Новосибирск: Изд-во ИМ СО РАН, 2009. С. 207–211.

**USE OF COMPUTER TECHNOLOGIES AND MATHEMATICAL METHODS  
FOR THE ANALYSIS OF AUTO-TRANSPORT NETWORKS**

**A.M. Purtov**

Associate Professor, Ph.D. (Eng.), e-mail: andr.purtov.yandex.ru

Branch of Sobolev Institute of Mathematics of Siberian Branch of Russian Academy of  
Sciences

**Abstract.** The way of application of taxonomy, reduction of graphs and geoinformation systems methods is developed for the analysis of routes in transport networks. The technology is shown on examples of the analysis of popular routes of Omsk. The example of use of taxonomy for the analysis of results of a reduction of graphs are shown. Similarity of results of visual and automatic taxonomy is shown. Results of taxonomy are represented on GIS — map of the graph of delays. The imitating model of crossroads is developed. The model is programmed on GPSSW.

**Keywords:** geoinformation system, simulation, reduction of graphs, taxonomy, analysis of data, auto transport networks.