



А.А. Бородинов

ОПРЕДЕЛЕНИЕ ПРЕДПОЧТЕНИЙ УЧАСТНИКА ДОРОЖНОГО ДВИЖЕНИЯ НА ЛИЧНОМ ТРАНСПОРТЕ МЕТОДОМ ПАРНЫХ СРАВНЕНИЙ

(Самарский университет)

В данной работе рассматривается проблема ранжирования объектов, используя информацию об их парном сравнении. В качестве «объектов» и «предпочтений» мы рассматриваем соответственно набор решений и выбор пользователя. Например, в задаче навигации - список маршрутов и маршрут, выбранный пользователем.

Методы парного сравнения первоначально использовались для ранжирования объектов, которые не могут быть описаны вектором признаков (например, организовывать виды фруктов для продажи в магазинах). Результаты таких сравнений обычно определяются матрицей (c_{ij}) [1,2]. Каждый элемент матрицы является абсолютной частотой i -го объекта над j -й. Для анализа таких данных в первой половине XX века были предложены две модели: модель Терстоуна [3] и модель Брэдли-Терри [4].

Предлагаемый метод может быть представлен как набор шагов:

- 1) нормализация значений признаков в диапазоне $[0,1]$;
- 2) выбор нового признака пространства (базиса) Y ;
- 3) преобразование исходного вектора признаков в новое пространство признаков Y с большей размерностью $K = \dim(Y) \geq N$;
- 4) построение линейного или нелинейного классификатора в пространстве признаков Y . Восстановление функции полезности (UF) для линейного классификатора и / или функции предпочтения (PF) выполняется автоматически;
- 5) оценка качества классификатора здания по набору данных испытаний;
- 6) если оценка удовлетворительная, прекратить процедуру; в противном случае перейдите к шагам 3 или 2 (если все доступные размеры пространства объектов уже используются).

Предложенный подход схематически изображен на рисунке 1 и обозначен пунктирной линией.

Работоспособность и эффективность предложенного метода были исследованы на задаче определения пользовательских предпочтений при планировании и осуществлении корреспонденций в городе на личном транспорте. В данной задаче пользователю предлагается несколько вариантов осуществления конкретной корреспонденции от точки выезда до точки прибытия, и задача метода заключается в том, чтобы сделать тот же выбор из предлагаемых вариантов корреспонденции, что и пользователь.

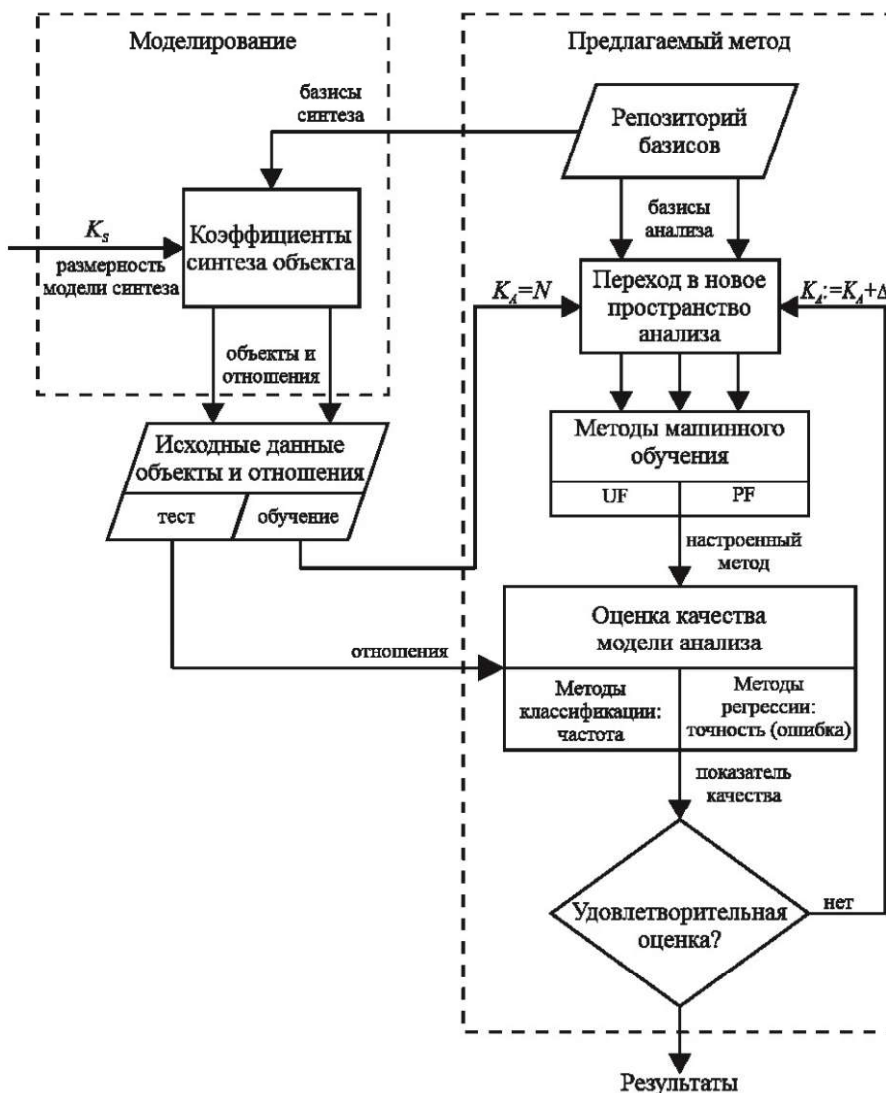


Рис. 1. Схема предлагаемого метода

Эксперимент ставился на данных, полученных от сотрудников кафедры геоинформатики и информационной безопасности Самарского университета, имеющих личный транспорт и ежедневно совершающих поездки по городу Самара на протяжении как минимум трех лет. В проведенном эксперименте каждому из четверых участвующих сотрудников предлагалось не менее 50 раз выбрать один из $S=5$ автоматически генерируемых путей от некоторой точки А до точки В.

Пример интерфейса работы с пользователем и варианты выбора представлены на рисунке 2. Каждый раз пары точек А и В менялись, повторы исключались. Таким образом, для каждого участника эксперимента формировалась общая выборка из как минимум 50 прецедентов пользовательских решений, которая в дальнейшем использовалась для формирования обучающих и контрольных выборок при кросс-валидации (число случайных разбиений при кросс-валидации выбиралось равным 5 и не зависело от пропорций деления выборки на обучающую и контрольную).



Всего анализировалось четыре варианта разбиения выборки, представленных в таблице 1. Первый из вариантов, при котором в обучающей выборке было всего три принятых пользователем решения, соответствовал ситуации "холодного старта" системы. Очевидно, при числе γ принятых решений количество обучающих пар составляло $|\Theta| = \gamma(S-1)$. Показателем качества работы метода выступала вероятности ошибок в реконструированных предпочтениях в парах.

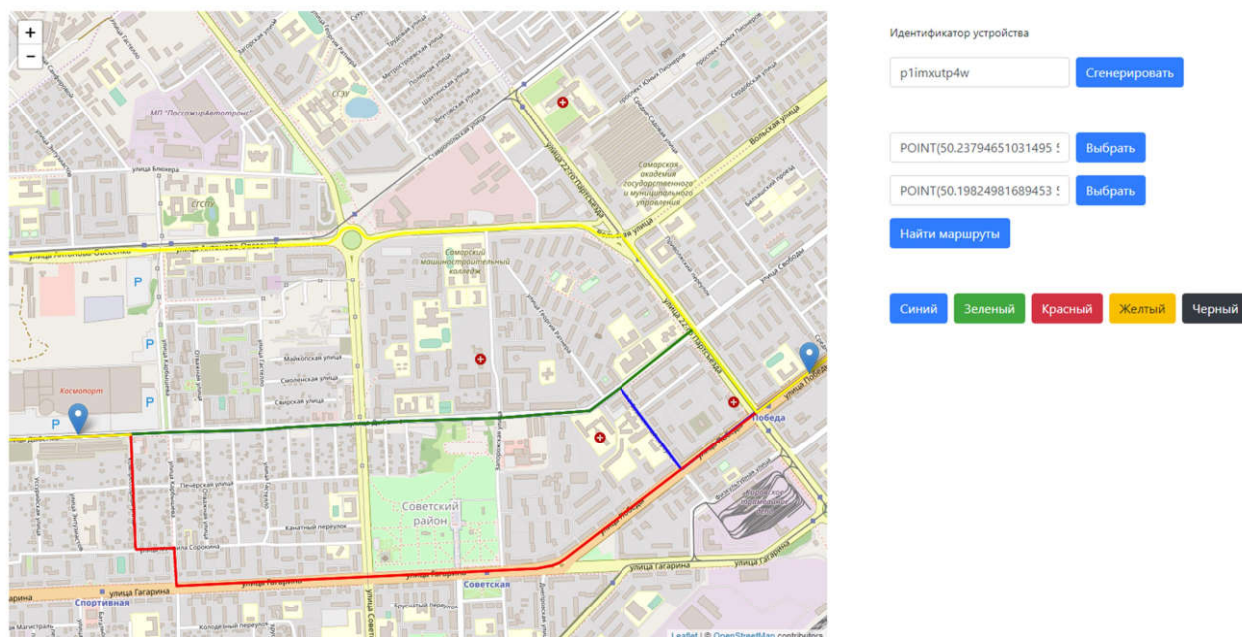


Рис. 2. Пример интерфейса работы с пользователем и варианты выбора

Таблица 1 – Параметры кросс-валидации

Размер обучающей выборки, в количестве принятых решений γ	Размер контрольной выборки, в количестве принятых решений
3	40
10	40
20	30
30	20

Для описания треков использовались следующие признаки:

- 1) отношение расстояния между А и В по прямой к длине текущего трека;
- 2) отношение длины кратчайшего (по расстоянию) трека к длине текущего трека;
- 3) интенсивность перекрестков на 100м текущего трека (=число перекрестков/(L/100));
- 4) отношение числа поворотов к общему числу перекрестков для текущего трека;



5) отношение числа перекрестков для трека с минимальным их количеством к числу перекрестков на текущем треке;

6) отношение числа левых поворотов на текущем треке к общему числу поворотов на нем;

7) отношение минимального времени движения (с максимально разрешенной скоростью) к оцененному времени движения по текущему треку;

8) отношение оцененного минимального времени движения по кратчайшему (по времени) треку к оцененному времени движения по текущему треку;

9) корень из дисперсии разброса текущих величин скорости на текущем треке по отношению к допустимой максимальной скорости (60 км/ч).

Результаты проведенных исследований представлены в таблице 2.

Таблица 2 – Зависимость вероятности ошибки \tilde{d} и размерности пространства анализа K_A от размеров обучающей выборки

№ user	$\gamma=3$		$\gamma=10$		$\gamma=20$		$\gamma=30$	
	\tilde{d}	K_A	\tilde{d}	K_A	\tilde{d}	K_A	\tilde{d}	K_A
1	0.205	11	0.198	13	0.200	15	0.181	13
2	0.225	17	0.205	13	0.194	11	0.168	13
3	0.197	15	0.190	13	0.159	13	0.188	13
4	0.245	13	0.235	11	0.172	11	0.219	11
среднее	0.218	14	0.207	12.5	0.181	12.5	0.189	12,5
медиана	0.215	14	0.202	13	0.183	12	0.185	13

Основываясь на полученных результатах можно сделать вывод, что предложенный метод является работоспособным - диапазон ошибок на всех рассматриваемых постановках задачи оказался в диапазоне 0.16-0.25 и качество конструируемого решения на реальных данных ожидаемо выше для больших размеров обучающей выборки.

Благодарности

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (уникальный идентификатор проекта RFMEFI57518X0177).

Литература

1. Fishburn PC. Utility theory for decision making. Wiley; 1970.
2. Fürnkranz J, Hüllermeier E, editors. Preference Learning. Berlin Heidelberg: Springer-Verlag; 2011. 10.1007/978-3-642-14125-6.
3. Thurstone LL. A law of comparative judgment. Psychological Review 1927; 34: 273–86. 10.1037/h0070288.



4. Bradley RA, Terry ME. Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika* 1952; 39: 324–45. 10.2307/2334029.

О.К. Головнин

ПРЕДСКАЗАТЕЛЬНАЯ МОДЕЛЬ ВЫЯВЛЕНИЯ ВОЗМОЖНЫХ УГРОЗ БЕЗОПАСНОСТИ ФУНКЦИОНИРОВАНИЯ ТРАНСПОРТНОЙ СИСТЕМЫ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ФЕДЕРАТИВНЫХ ДАННЫХ SMART CITY

(Самарский университет)

Нештатные ситуации, влияющие на нормальное функционирование транспортной системы, остаются ощутимой проблемой как для городских поселений и агломераций, так и для удаленных транспортных районов [1, 2]. Возникновение и развитие нештатной ситуации может привести к нарушению связности территорий, что оказывает непосредственное влияние на безопасность транспортной системы в целом [3, 4]. Развитие интегрирующих систем «Умных городов» (Smart City) позволяет выработать новые решения, направленные на снижение влияния дестабилизирующих факторов на функционирование транспортной системы [5-7].

В работе представлена предсказательная модель и программный комплекс на основе интеллектуального (нейросетевого) анализа наборов геопривязанных данных, порождаемых интегрированными средствами Smart City. Предлагаемые решения предназначены для повышения эффективности выявления возможных угроз безопасности функционирования транспортной системы, при этом рассматриваются только транспортные системы, образованные на основе улично-дорожной сети общего пользования, в пределах которой осуществляется дорожное движение.

Проведено исследование структур данных, используемых в многочисленных системах класса Smart City, которое позволило сформировать требования к характеристикам входных и выходных слоев искусственной нейронной сети, наборам данных для обучения. Данные, поступающие из различных подсистем Smart City, как правило, обладают геопространственной меткой и в совокупности имеют большой объем. Таким образом, обработка наборов данных строится на основе применения геоинформационных технологий и алгоритмов обработки больших данных.

Проведены исследования по формированию наиболее эффективной внутренней структуры используемой нейронной сети, а также выполнена калибровка механизмов предобработки информации.

Разработана предсказательная модель для выявления возможных угроз безопасности функционирования транспортной системы, построенная на основе искусственной нейронной сети с предобработкой больших