

УДК 656.13

*С. А. АЗЕМША, кандидат технических наук, Белорусский государственный университет транспорта, г. Гомель*

## ВЫБОР ДИСКРИМИНИРУЮЩИХ ПЕРЕМЕННЫХ В ОТНОШЕНИИ ОСТАНОВОЧНЫХ ПУНКТОВ МАРШРУТНЫХ ТРАНСПОРТНЫХ СРЕДСТВ

Стратегическая важность пассажирских перевозок обуславливает необходимость уделять этому виду перевозок значительное внимание. От качества удовлетворения спроса на городские пассажирские перевозки во многом зависит успешное экономическое развитие города или региона. Общеизвестно, что одним из важнейших показателей качества пассажирских перевозок является скорость доставки пассажиров. Ее величина зависит от целого ряда факторов: интенсивности движения транспортных средств, вида и состояния дорожного покрытия, метеорологических условий, организации дорожного движения и др. Характерной особенностью регулярных городских перевозок пассажиров является наличие значительного числа остановочных пунктов, расположенных друг от друга на небольшом удалении. Этот факт оказывает ощутимое влияние на скорость сообщения и, следовательно, на общую скорость доставки пассажиров. Кроме того, при простоях под посадкой-высадкой маршрутные транспортные средства продолжают потреблять энергию (электрическую или углеводородное топливо), не выполняя при этом транспортной работы.

Приведена методика расчета параметров, влияющих на различия остановочных пунктов, и определяются такие параметры.

**М**ассовость городских перевозок пассажиров маршрутными транспортными средствами регулярного сообщения обуславливает необходимость постоянного контроля за качеством данного вида услуг. При этом необходимо уделять достаточное внимание каждому элементу транспортного процесса, в том числе и продолжительности простоев на остановочных пунктах.

В настоящее время порядок посадки-высадки пассажиров маршрутных транспортных средств регламентирован п. 222 Правил автомобильных перевозок пассажиров. Из указанного пункта следует, что водитель маршрутного транспортного средства должен останавливаться для посадки-высадки пассажиров как можно ближе к передней границе остановочного пункта [1]. Однако это требование далеко не всегда соблюдается. Кроме того, в нормативных актах не регламентируется время простоя маршрутного транспортного средства в зоне остановочного пункта. Негативные результаты этого ощущает на себе каждый пассажир:

1 Водители маршрутных транспортных средств зачастую без каких-либо видимых причин простаивают на остановочных пунктах после прекращения посадки-высадки пассажиров.

2 Нередки простои маршрутных транспортных средств в ожидании освобождения остановочного пункта вследствие его занятости.

3 Неправильное расположение маршрутных транспортных средств на остановочных пунктах приводит к неоправданным простоям следующих за ними маршрутных транспортных средств и т. д.

Требования к информационному обеспечению остановочных пунктов маршрутных транспортных средств и к перевозке пассажиров сформулированы в СТБ [2] и [3] соответственно. Содержащиеся в упомянутых нормативных актах положения едины для всех остановочных пунктов, вне зависимости от пассажирообмена, количества маршрутов, проходящих через остановочный пункт, и других параметров.

Установлено, что общая задержка каждого маршрутного транспортного средства на остановочном пункте в среднем составляет 20 с [4, 5]. При этом суммарное время простоя маршрутного транспортного средства на остановочном пункте

$$t_6 = t_1 + t_2 + t_3 + t_4 + t_5, \quad (1)$$

где  $t_1$  – простои перед остановочным пунктом вследствие его занятости, с;  $t_2$  – простои перед остановочным пунктом вследствие неправильного расположения на нем впереди идущих маршрутных транспортных средств, с;  $t_3$  – простои на остановочном пункте при посадке-высадке пассажиров, с;  $t_4$  – простои на остановочном пункте без посадки-высадки пассажиров с открытыми дверями, с;  $t_5$  – простои на остановочном пункте с закрытыми дверями в ожидании освобождения проезда, с.

В результате исследований установлено [4–6], что в среднем за сутки автобусы в г. Гомеле совершают 71,7 ч непроизводительных простоев. При этом они расходуют топлива более чем на 215,3 млн руб. в год. Микроавтобусы в среднем за сутки в г. Гомеле совершают 12,7 ч непроизводительных простоев. При этом они расходуют топлива более чем на

6,9 млн руб. в год. Троллейбусы г. Гомеля совершают в среднем 142,3 ч непроизводительных простоев в сутки. При этом они потребляют электроэнергии более чем на 87,4 млн руб. за год. С учетом стоимости потерянного пассажирами времени совокупные потери от непроизводительных простоев маршрутных транспортных средств в зоне остановочных пунктов г. Гомеля в будние дни составляют более 16,1 млрд руб. за год [6].

Вышеприведенные доводы формируют предпосылки к исследованию остановочных пунктов маршрутных транспортных средств с целью их классификации и, на основании этого, обоснования параметров, влияющих на качество обслуживания пассажиров. Классификации какого-либо объекта должен предшествовать анализ дискриминантных функций, который используется для принятия решения о том, какие переменные различают (дискриминируют) две или более возникающие совокупности (группы). Основная идея анализа дискриминантных функций заключается в том, чтобы определить, отличаются ли совокупности по среднему какой-либо переменной, а затем использовать эту переменную, чтобы предсказать для новых членов принадлежность к той или иной совокупности.

Исследуемая выборка представлена девятью остановочными пунктами маршрутных транспортных средств г. Гомеля. Зависимый параметр – наименование остановочного пункта. Независимыми переменными являются:

- количество вошедших / вышедших пассажиров ( $A_{sum}$ );
- простои маршрутных транспортных средств перед остановочным пунктом вследствие его занятости ( $t_1$ );
- простои маршрутных транспортных средств перед остановочным пунктом вследствие неправильного расположения других маршрутных транспортных средств на остановочном пункте ( $t_2$ );
- простои маршрутных транспортных средств на остановочном пункте с открытыми дверями без посадки-высадки пассажиров ( $t_4$ );
- простои маршрутных транспортных средств на остановочном пункте в ожидании начала движения впереди остановившихся маршрутных транспортных средств ( $t_5$ );
- суммарные простои маршрутных транспортных средств перед и на остановочном пункте с учетом времени на посадку-высадку пассажиров ( $t_6$ );
- общее число маршрутов маршрутных транспортных средств, проходящих через данный остановочный пункт ( $N_s$ ).

Общий объем выборки для наблюдаемой совокупности параметров составил 1903 значения.

Дискриминантный анализ выполнялся с помощью программы *STATISTICA* методом пошагового последовательного включения: на каждом шаге система *STATISTICA* просматривает все переменные и находит, какая из них вносит наибольший вклад в различие между совокупностями. Эта переменная в таком случае должна быть включена в модель, и на данном шаге система *STATISTICA* переходит к следующему шагу. Результаты первого этапа анализа приведены на рисунке 1.

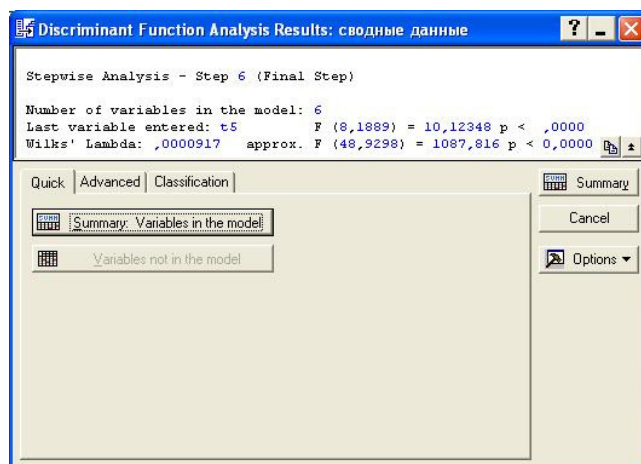


Рисунок 1 – Предварительные результаты анализа

Из рисунка 1 видно, что анализ остановлен на шестом шаге (Step6), число переменных, включенных в модель – шесть, последней в модель была добавлена переменная  $t_5$ . Статистика Уилкса лямбда (Wilks' Lambda) равна 0,0000917. Статистика лямбда Уилкса для общего разделения вычисляется, как отношение детерминанта матрицы внутригрупповых дисперсий/ковариаций к детерминанту общей ковариационной матрицы. В общем, статистика Уилкса лямбда является стандартной статистикой, используемой для обозначения статистической значимости мощности дискриминации в текущей модели. Ее значение меняется от 1,0 (нет никакой дискриминации) до 0,0 (полная дискриминация). Значение данной статистики говорит о высокой степени дискриминации в рассматриваемой модели. Критерий Фишера ( $F$ ) и уровень статистической значимости ( $p$ ) говорят о высокой степени значимости полученной модели.

На рисунке 2 представлена таблица независимых вкладов каждой из шести переменных в общую дискриминацию между остановочными пунктами. Кроме прочих характеристик, приведена частная лямбда Уилкса (Partial Lambda) – статистика Уилкса лямбда для одиночного вклада соответствующей переменной в дискриминацию между совокупностями. Это значение можно рассматривать как аналог частного коэффициента корреляции. Частная лямбда вычисляется как мультипликативное приращение в лямбде, которое является результатом добавления соответствующей переменной

ной. Иначе, частная *лямбда* есть отношение *лямбды* Уилкса после добавления соответствующей переменной к *лямбде* Уилкса до добавления этой переменной. Так как *лямбда* с величиной 0,0 обозначает полную дискриминацию, то чем ниже ее значение в этом столбце, тем больше одиночный вклад соответствующей переменной в степень дискриминации. На данном этапе исследования видно, что наибольший вклад в дискриминацию остановочных пунктов вносит переменная  $N_s$  (общее число маршрутов, проходящих через остановочный пункт), а наименьший –  $t_5$  (простой маршрутных транспортных средств перед остановочным пунктом вследствие его занятости).

Discriminant Function Analysis Summary (сводные данные)						
Step 6, N of vars in model: 6; Grouping: Ost (9 grps)						
Wilks' Lambda: .00009 approx. F (48,9298)=1087,8 p<0,0000						
N=1903	Wilks' Lambda	Partial Lambda	F-remove	p-level	Toler.	1-Toler. (R-Sqr.)
$N_s$	0,540436	0,000170	1390787	0,000000	0,999491	0,000509
Asum	0,000109	0,844686	43	0,000000	0,640191	0,359809
t2	0,000101	0,905357	25	0,000000	0,664702	0,335298
t4	0,000103	0,888649	30	0,000000	0,658616	0,341384
t6	0,000098	0,932005	17	0,000000	0,402217	0,597783
t5	0,000096	0,958889	10	0,000000	0,925838	0,074162

Рисунок 2 – Независимый вклад переменных в общую дискриминацию остановочных пунктов

Статистика Уилкса *лямбда* может быть преобразована к стандартному  $F$  значению, и можно вычислить соответствующее  $p$ -значение для каждого значения  $F$ . Значение  $F$ -статистики определяется следующим образом:

$$F = \left( \frac{n - q - p}{q - 1} \right) \left( \frac{1 - \lambda_c}{\lambda_c} \right), \quad (2)$$

где  $n$  – количество наблюдений;  $q$  – количество групп;  $p$  – количество переменных;  $\lambda_c$  – частная *лямбда*.

Из третьего столбца таблицы, изображенной на рисунке 2, видно, что значения  $F$ -статистики для каждой из шести переменных, включенных в модель, больше минимально допустимого значения ( $F\text{-remove} = 8,1889$ ).

Чтобы избежать плохой обусловленности матриц, система *STATISTICA* должна постоянно проверять так называемые значения *толерантности* (Toler) для каждой переменной. Эти значения также неизменно отображаются, когда вы запрашиваете программу выдать значения статистик, включенных в модель, и тех, которые в модель не включены. Значение *толерантности* вычисляется как  $1 - R\text{-квадрат}$ , где  $R\text{-квадрат}$  – коэффициент множественной корреляции для соответствующей переменной со всеми другими переменными в текущей модели. Таким образом, это есть доля дисперсии, относящаяся к соответствующей переменной. В общем, когда переменная почти полностью избыточна (и поэтому матрица задачи является

плохо обусловленной), значение *толерантности* для этой переменной будет приближаться к нулю. Значением по умолчанию в модуле *Дискриминантный анализ* для минимально приемлемой *толерантности* является 0,01. Система *STATISTICA* выдает сообщение о плохо обусловленной матрице, когда *толерантность* для какой-либо переменной падает ниже этого значения, т. е. если всякая переменная более чем на 99 % сократима (пользователь может изменить это значение).

Чтобы увидеть, как шесть переменных разделяют различные совокупности (остановочные пункты), можно вычислить действительную дискриминантную функцию. Для этого необходимо определить, являются ли дискриминантные функции (корни) статистически значимыми (рисунок 3).

Chi-Square Tests with Successive Roots Removed (сводные данные)						
Roots Removed	Eigenvalue	Canonical R	Wilks' Lambda	Chi-Sqr.	df	p-level
0	6309,933	0,999921	0,000092	17612,35	48	0,000000
1	0,294	0,476533	0,578956	1035,40	35	0,000000
2	0,157	0,368026	0,749054	547,40	24	0,000000
3	0,108	0,312837	0,866402	271,68	15	0,000000
4	0,037	0,188232	0,960393	76,56	8	0,000000
5	0,004	0,065792	0,995671	8,22	3	0,041708

Рисунок 3 – Результат оценки статистической значимости дискриминантных функций

В общем, эта таблица дает отчет о пошаговом критерии с включением для всех канонических корней. Первая строка всегда содержит критерий значимости для всех корней, вторая строка дает данные о значимости корней, оставшихся после удаления первого корня, и т. д. Таким образом, эта таблица показывает, как много канонических корней (дискриминирующих функций) следует интерпретировать. Очевидно, что статистически значимы все пять дискриминирующих переменных.

Далее необходимо определить стандартизованные коэффициенты дискриминантной функции (рисунок 4). Стандартизованные коэффициенты – это те коэффициенты, которые обычно используются для интерпретации, так как они относятся к нормированным переменным и поэтому должны находиться в сравнимых масштабах.

Standardized Coefficients (сводные данные) for Canonical Variables						
Variable	Root 1	Root 2	Root 3	Root 4	Root 5	Root 6
$N_s$	-1,000	0,001028	-0,000273	0,00293	0,00153	0,000669
Asum	0,018	-0,958282	0,269998	-0,41056	0,42593	-0,469607
t2	-0,003	0,071565	0,052538	-1,18891	0,26240	-0,119032
t4	-0,011	-0,049752	1,032065	-0,47865	0,15453	0,444528
t6	0,005	-0,005983	-0,816150	0,58232	-1,12260	0,469808
t5	0,001	-0,183649	-0,301043	-0,03153	0,80667	0,551390
Eigenval	6309,933	0,293801	0,156662	0,10848	0,03673	0,004347
Cum.Prop	1,000	0,999951	0,999976	0,99999	1,00000	1,00000

Рисунок 4 – Стандартизованные коэффициенты дискриминантной функции

Из рисунка 4 видно, что первая дискриминантная функция (Root1) взвешивается наиболее тяже-

ло количеством маршрутов, проходящих через остановочный пункт ( $N_s$ ), вторая – переменной  $A_{sum}$ , третья – переменными  $t_4$  и  $t_6$ , четвертая – переменной  $t_2$ , пятая – переменными  $t_6$  и  $t_5$ , шестая – переменной  $t_5$ .

На рисунке 4 приведены также собственные значения (корни) для каждой дискриминантной функции (Eigenval) и кумулятивная доля объясненной дисперсии (Cum. Prop), накопленной каждой функцией.

Следующим этапом необходимо определить коэффициенты факторной структуры (рисунок 5). Коэффициенты факторной структуры представляют корреляцию между переменными дискриминирующей функцией и в общем случае используются для интерпретации "значимости" дискриминирующей функции.

Factor Structure Matrix (сводные данные) Correlations Variables - Canonical Roots (Pooled-within-groups correlations)						
Variable	Root 1	Root 2	Root 3	Root 4	Root 5	Root 6
Ns	-0,999771	-0,018778	-0,007567	0,001405	-0,001320	-0,006609
Asum	-0,001473	-0,977850	0,010794	-0,065119	-0,106567	-0,167610
t2	-0,002526	0,181264	-0,392229	-0,886549	-0,162051	0,032607
t4	-0,000656	-0,207615	0,662248	-0,106600	-0,257919	0,663655
t6	-0,001638	-0,500713	-0,269384	-0,224673	-0,586473	0,531308
t5	0,001337	-0,199654	-0,379846	0,030524	0,581891	0,690163

Рисунок 5 – Коэффициенты факторной структуры

Теперь, зная, какие переменные участвуют в дискриминации между различными остановочными пунктами, можно определить природу дискриминации для каждого канонического корня. В качестве первого шага рассмотрим канонические средние (рисунок 6).

Group	Means of Canonical Variables (сводные данные)					
	Root 1	Root 2	Root 3	Root 4	Root 5	Root 6
Универмаг "Гомель"-Цирк	-16,4169	-1,06306	-0,231953	0,117081	0,209070	0,021199
Универмаг "Гомель"-Университет	-16,4147	-1,30049	0,138867	-0,187951	-0,351036	-0,014553
8 Марта-Универмаг "Гомель"	-27,4110	0,28019	-0,049189	0,391194	-0,163007	-0,007801
Цирк-ул. Жарковского	49,3612	0,14556	1,024052	-0,205734	0,075520	0,117397
ул. Головацкого-Универсам	126,8317	0,03712	0,697211	-0,255459	-0,309798	0,083098
ул. Огоренко-Ледовый дворец	170,0259	0,23338	-0,481664	0,047849	0,113539	0,024323
Коминтерн-ул. Гагарина	-71,2879	0,25312	0,240645	0,027641	0,248482	-0,016440
ЗИП-ул. Барыкина	-60,3170	0,35334	-0,401939	-0,629863	-0,079607	0,006226
ул. Головацкого-Дружба	137,1197	-0,07445	0,597333	-0,300050	0,035568	-0,313299

Рисунок 6 – Канонические средние переменных

Очевидно, что первая дискриминантная функция отделяет главным образом остановочные пункты «ул. Головацкого-Универсам», «ул. Огоренко-Ледовый дворец» и «ул. Головацкого-Дружба» от других остановочных пунктов. Каноническое среднее этих остановочных пунктов очень сильно отличается от канонических средних других остановочных пунктов. Вторая дискриминантная функция предназначена для разделения главным образом остановочных пунктов «ЗИП-ул. Барыкина», «Коминтерн-ул. Гагарина» и «8 Марта-Универмаг «Гомель» от других остановочных пунктов. Однако, основываясь на рассмотренных ранее собственных значениях (рисунок 4, строка «Eigenval»), качество дискриминации этой функции хуже.

Для оценки качества классификационной модели можно построить матрицу классификации (рисунок 7). В первой строке приведены априорные вероятности для каждой совокупности. В первом столбце таблицы приведен процент наблюдений, которые были правильно классифицированы для каждой совокупности полученными функциями классификации. Оставшиеся столбцы дают число случаев правильной и неправильной классификации для каждой совокупности.

Group	Classification Matrix (сводные данные) Rows: Observed classifications Columns: Predicted classifications								
	Percent Correct	Универмаг "Гомель"-Цирк p=,11686	Универмаг "Гомель"-Университет p=,06096	8 Марта-Универмаг "Гомель" p=,26852	Цирк-ул. Жарковского p=,05780	ул. Головацкого-Универсам p=,03363	ул. Огоренко-Ледовый дворец p=,11613	Коминтерн-ул. Гагарина p=,17551	ЗИП-ул. Барыкина p=,1392
Универмаг "Гомель"-Цирк	88,7387	197	25	0	0	0	0	0	
Универмаг "Гомель"-Университет	33,6207	77	39	0	0	0	0	0	
8 Марта-Универмаг "Гомель"	100,0000	0	0	511	0	0	0	0	
Цирк-ул. Жарковского	100,0000	0	0	0	110	0	0	0	
ул. Головацкого-Универсам	98,4375	0	0	0	0	63	1	0	
ул. Огоренко-Ледовый дворец	100,0000	0	0	0	0	0	221	0	
Коминтерн-ул. Гагарина	100,0000	0	0	0	0	0	0	334	
ЗИП-ул. Барыкина	100,0000	0	0	0	0	0	0	0	
ул. Головацкого-Дружба	100,0000	0	0	0	0	0	0	0	
Total	94,5875	274	64	511	110	63	222	334	

Рисунок 7 – Матрица классификации

**Выводы.** В данной статье обосновывается целесообразность проведения мероприятий по изучению процессов, происходящих на остановочных пунктах маршрутных транспортных средств. С целью снижения непроизводительных простоев маршрутных транспортных средств предложено произвести классификацию остановочных пунктов. Для этого необходимо определить параметры, различающие совокупности остановочных пунктов. Установлено, что такими переменными являются:

- общее число маршрутов маршрутных транспортных средств, проходящих через данный остановочный пункт ( $N_s$ );
- количество вошедших – вышедших пассажиров ( $A_{sum}$ );
- простой маршрутных транспортных средств перед остановочным пунктом вследствие неправильного расположения других маршрутных транспортных средств на остановочном пункте ( $t_2$ );

– простои маршрутных транспортных средств на остановочном пункте с открытыми дверями без посадки/высадки пассажиров ( $t_4$ );

– суммарные простои маршрутных транспортных средств перед и на остановочном пункте с учетом времени на посадку-высадку пассажиров ( $t_6$ );

– простои маршрутных транспортных средств на остановочном пункте в ожидании начала движения впереди остановившихся маршрутных транспортных средств ( $t_5$ ).

Статистический анализ функций дискриминации и дискриминационных переменных показывает статистическую значимость полученных моделей. При этом функции классификации правильно классифицируют объект (остановочный пункт) в 94,5 % случаев.

#### Список литературы

1 О некоторых вопросах автомобильных перевозок пассажиров: постановление Совета Министров Респ. Беларусь 30 июня 2008 г. № 972 // Эталон – Беларусь [Электронный ресурс] / Нац. центр правовой информ. Респ. Беларусь. – Минск, 2009.

Получено 22.04.2011

#### **S. A. Azemsha.** Choice of discriminating variables concerning stopping points of routing vehicles.

Strategic importance of passenger transportations causes necessity to pay to this kind of transportations considerable attention. The successful economic development of the city in many respects depends on quality of satisfaction of demand for city passenger transportations or region. Well-known that one of the major indicators of quality of passenger transportations is speed of delivery of passengers. Its size depends on variety of factors: intensity of movement of vehicles, a kind and a condition of a road covering, weather conditions, the organization of traffic, etc. Prominent feature of regular city transportations of passengers presence of considerable number of the stopping points located from each other on small removal is. This fact makes notable impact on speed of the message and, hence, on the general speed of delivery of passengers. Besides, at idle times under landing-disembarkation routing vehicles continue to consume energy (electric or hydrocarbonic fuel), without performing thus transport work.

In article the technique of definition of the parameters influencing distinctions of stopping points is resulted, and such parameters are defined.

2 Сооружения станционные пассажирские, автомобили, троллейбусы, трамваи: СТБ 1389-2003, с изм. и доп. – Введ. 2003-03-19. – Минск : Межгос. совет по стандартизации, метрологии и сертификации; Беларус. гос. ин-т стандартизации и сертификации, 2003. – 33 с.

3 Услуги по перевозке пассажиров автомобильным транспортом. Ч. 2. Требования к перевозке автобусами в регулярном сообщении: СТБ 1731.2-2007. – Введ. 2007-02-23. – Минск : Межгос. совет по стандартизации, метрологии и сертификации; Беларус. гос. ин-т стандартизации и сертификации, 2003. – 5 с.

4 Снижение задержек транспортных средств на остановочных пунктах при городских перевозках пассажиров в регулярном сообщении / С. А. Аземша [и др.] // Вестник Белорус. гос. ун-та трансп. : Наука и транспорт. – 2009. – № 2. – С. 21–24.

5 **Аземша, С. А.** Социально-экономическая оценка временных потерь пассажиров маршрутных транспортных средств в городском регулярном сообщении / С. А. Аземша, С. В. Скирковский, В. Н. Стукачев // Совершенствование организации дорожного движения и перевозок пассажиров и грузов : материалы междунар. науч.-практ. конф.; редкол.: Ф. А. Романюк [и др.]. – Минск : БНТУ, 2010. – С. 64–69.

6 **Аземша, С. А.** Экономическая оценка непроизводительных простоев маршрутных транспортных средств в зоне остановочных пунктов при выполнении регулярных пассажирских перевозок в г. Гомеле / С. А. Аземша, В. Н. Стукачев // Вестник Белорус. гос. ун. трансп. : Наука и транспорт. – 2010. – № 2 (21). – С. 30–37.