DOI dx.doi.org/10.24866/VVSU/2073-3984/2019-2/058-069

УДК 656.072:519.211

В.Н. Ембулаев

Владивостокский государственный университет экономики и сервиса
Владивосток. Россия

Вероятностный метод определения поездок пассажиров на маршруте по данным входа и выхода

При разработке расчётного способа определения поездок пассажиров по данным входа и выхода было отмечено, что число корреспондирующих пассажиров между двумя конкретными остановками на маршруте представляет собой дискретную случайную величину. Для каждого значения этой случайной величины вычисляется вероятность. Установление соответствия между численными значениями дискретной случайной величины и их вероятностями позволило сделать вывод, что она подчиняется данному закону распределения. И тогда в качестве определения поездки между двумя конкретными остановками на маршруте принимается то значение случайной величины, вероятность которого наибольшая.

Ключевые слова и словосочетания: пассажиропотоки, случайная величина, вероятность события, закон распределения, наивероятнейшее число.

V.N. Embulaev

Vladivostok State University of Economics and Service
Vladivostok. Russia

A probabilistic method of determining a travel route according to the input and output

In developing a calculation method for determining passenger travel by input and output data, it was noted that the number of corresponding passengers between two specific stops on the route is a discrete random variable. For each value of this random variable, the probability is calculated. Establishing a correspondence between the numerical values of a discrete random variable and their probabilities allowed us to assume that it is subject to this distribution law. And then, as a definition of the trip between two specific stops on the route, the value of the random variable, the probability of which is the greatest, is taken.

Keywords: passengers flows, casual magnitude, probability of event, law of distribution, the most probable number.

Ембулаев Владимир Николаевич – д-р экон. наук, профессор; профессор кафедры математики и моделирования ВГУЭС; e-mail: Vladimir.Embulaev@vvsu.ru

Введение. Транспортная система – особый сектор экономики крупного города. В сравнении с другими секторами городской экономики результаты её деятельности так или иначе ощущают на себе все жители. Более того, сохранение социальной, экономической и политической стабильности жизни города во многом зависит именно от эффективности работы транспортной системы.

В современных крупных городах социально-экономическое развитие отличается тем, что потребности населения в поездках растут опережающими темпами по сравнению с ростом возможностей транспортных систем в их удовлетворении. Поэтому организация и управление перевозками пассажиров по маршрутам крупных городов являются трудно разрешимыми задачами [2, 7]. Сложность решения этих задач заключается в том, что каждый раз приходится не заново их решать, а корректировать уже имеющиеся решения на основе поступления новой информации о потребностях населения в поездках. Однако специалисты считают, что эффективное решение этих задач возможно лишь в условиях наиболее рационального использования имеющихся в наличии и готовых к эксплуатации транспортных средств (ТС) на основе разработки и внедрения автоматизированных систем управления всеми видами городского пассажирского транспорта [5; 6].

В связи с тем, что основной эффект от автоматизации управления достигается за счёт автоматизации информационных процессов, то исследование в области механизации и автоматизации получения информации о пассажиропотоках с использованием математических методов и ЭВМ является актуальным.

Постановка задачи. В городе ежедневно наблюдаются передвижения десятков, сотен тысяч людей с одного места в другое. И порой кажется, что эти передвижения являются хаотическими. Однако это совсем не так. Любое передвижение в городе для конкретного человека чем-то обусловлено. Но отсутствие такой информации о каждом передвигающемся человеке в городе вполне допускает считать эти передвижения случайными.

Кроме того, когда пассажир выбирает поездку по маршруту (i, j), где i и j – пункты отправления и назначения, то в процессе решения задач организации и управления транспортным обслуживанием населения в поездках по маршруту вовсе необязательно знать, почему он выбрал именно этот путь, а не другой. Соответственно и выбор пути (i, j) каждым пассажиром можно считать происходит случайно и независимо от подобных выборов другими пассажирами. Из этого следует, что изучение передвижений пассажиров по маршрутам города требует вероятностной интерпретации.

Известно, что в процессе исследования возможностей автоматизации и механизации получения информации о пассажиропотоках наибольшие результаты были достигнуты в области применения технических средств для фиксации данных входа (a_i , $1 \leqslant i \leqslant n$) и выхода ($b_j$, $1 \leqslant j \leqslant n$) пассажиров на каждом остановочном пункте (ОП) маршрута (где n – число ОП на маршруте) [1]. Однако анализ процессов обработки данных входа-выхода показал, что к числу малоисследованных задач относится задача получения информации о маршрутных корреспонденциях пассажиров (x_{ij} , $i \leqslant j$), в связи с чем и было указано на актуальность постановки и решения такой задачи.

В процессе решения задач организации и управления перевозками пассажиров по маршрутам крупного города большое значение имеет наличие информации о поездках пассажиров по маршруту (x_{ij}). Отсутствие доступных способов получения такой информации поставило задачу разработки расчётного способа получения x_{ij} в результате обработки данных входа и выхода пассажиров, фиксированных на каждом ОП маршрута [4].

Когда ТС стояло на j-м ОП маршрута, то в салоне находилось, допустим, Q_{j-1} пассажиров, которые приехали к данному ОП в результате отправления от предыдущего (j-1)-го ОП. После того, как произошёл обмен пассажирами, ТС отправилось к следующему (j+1)-му ОП, и в салоне уже находилось Q_j пассажиров, число которых определяется по формуле:

$$Q_j = \left(Q_{j-1} - b_j\right) + a_j = \sum_{r=1}^{j} \left(a_r - b_r\right). \qquad (1)$$

Во время сбора данных входа-выхода зафиксировали, что на i-м ОП в салон ТС вошло a_i пассажиров. Но когда ТС подошло к следующему (i+1)-му ОП, то некоторые из них могли и выйти, – это и будут как раз те пассажиры, которые совершили поездку от i-го до (i+1)-го ОП, которое обозначим через $x_{i,i+1}$. Вычитая это число из всех вошедших на i-м ОП, получим оставшихся пассажиров, которые продолжат свои поездки дальше по маршруту. Обозначим это число через $a_{i,i+1}$., которое вычисляется следующим образом: $a_{i,i+1} = a_i - x_{i,i+1}$.

Если рассматривать любые i и j на маршруте, где $i<j$, то данная величина вычисляется по формуле:

$$a_{ij} = a_i - \sum_{r=i+1}^{j-1} x_{ir} , \qquad (2)$$

причём для $j=i$ $a_{ij} = a_i$.

Другими словами, a_{ij} – это число оставшихся пассажиров из всех вошедших на i-м ОП, которое определяется вычитанием из a_i всех уже совершивших следующие поездки на маршруте: от i-го до (i+1)-го – $x_{i,i+1}$, от i-го до (i+2)-го – $x_{i,i+2}$,..., от i-го до (j-1)-го – $x_{i,j-1}$.

В перевозочном процессе по маршруту, когда на j-м ОП стояло ТС, то внутри салона находилось Q_{j-1} пассажиров, среди которых имелись и вошедшие на i-м ОП – a_{ij} . В результате обмена пассажирами на j-м ОП вместе с b_j могли выйти и те пассажиры, которые вошли в салон ТС на i-м ОП, т.е. из группы a_{ij} . В этом случае число x_{ij} , которое одновременно будет принадлежать a_{ij} и b_j , является искомой величиной.

Если таким образом определить все поездки пассажиров между двумя любыми ОП на маршруте, то их можно отобразить в виде таблицы элементов маршрутных корреспонденций пассажиропотоков (табл. 1).

Таблица 1

Матрица элементов маршрутных корреспонденций пассажиропотоков

Номера ОП входа	Номера ОП выхода							Вошло в ТС
	1	2	3	4	5	…	*n*	
1	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	…	x_{1n}	a_1
2		x_{22}	x_{23}	x_{24}	x_{25}	…	x_{2n}	a_2
3			x_{33}	x_{34}	x_{35}	…	x_{3n}	a_3
4				x_{44}	x_{45}	…	x_{4n}	a_4
5					x_{55}	…	x_{5n}	a_5
						…		
						…		
						…		
n							x_{nn}	a_n
Вышло	b_1	b_2	b_3	b_4	b_5	…	b_n	

Примечание: n – число ОП на маршруте; a_i – число пассажиров, вошедших в ТС на i-м ОП; b_j – число пассажиров, вышедших из ТС на j-м ОП; x_{ij} – число пассажиров, совершивших поездки от i-го до j-го ОП, $i \leq j$.

Из таблицы 1 видно, что математическая формализация постановки задачи определения x_{ij} по данным a_i и b_j заключается в следующем.

Задана система линейных алгебраических уравнений:

$$\sum_{j=i}^{n} x_{ij} = a_i \ ; \ \sum_{i=1}^{j} x_{ij} = b_j \ ; \ x_{ij} \geqslant 0; \ 1 \leqslant i \leqslant n, \ 1 \leqslant j \leqslant n, \qquad (3)$$

причём выполняется условие

$$\sum_{i=1}^{n} a_i = \sum_{j=1}^{n} b_j \ \ (\text{или} \ \sum_{i=1}^{n}\sum_{j=i}^{n} x_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{j} x_{ij} \ ).. \qquad (4)$$

Но система (3) состоит из 2n уравнений с $\frac{n(n+1)}{2}$ неизвестными, а такая система имеет единственное решение, когда число уравнений в системе равно числу неизвестных, т.е. единственное решение возможно только при выполнении условия: $2n(n+1)/2$. Следовательно, единственное решение возможно при $n > 3$, в то время как для n_3 существует, в принципе, бесчисленное множество решений, и потому без дополнительного предположения относительно распределения пассажирских потоков между ОП маршрута нельзя получить однозначного решения данной задачи.

Вероятностный метод решения задачи. При описании и постановке задачи определения поездок пассажиров на маршруте по данным входа и выхода можно принять следующее допущение: во время стоянки ТС на ОП маршрута для каждого пассажира в салоне событие выйти на этом ОП или поехать дальше считать равновероятным. И так как искомая величина x_{ij} должна одновременно принадлежать a_{ij} и b_j, то при $a_{ij} \geqslant b_j$ она может принимать целочисленные значения от 0 до b_j, а при $a_{ij} \leqslant b_j$ – от 0 до a_{ij}. В общем случае она может принимать значения не меньше 0 и не больше $min \lfloor a_{ij}, b_j \rfloor$.

Для того чтобы решить задачу (3–4), вначале осуществим постановку и решение следующей задачи с учётом принятого выше допущения.

Имеются данные входа (a_i, $1 \leqslant i \leqslant n$) и выхода ($b_j$, $1 \leqslant j \leqslant n$). На их основе, используя формулы (1) и (2), для двух конкретных ОП на маршруте i и j можно вычислить значения Q_{j-1} и a_{ij}. Надо найти вероятность того, что между *i* и *j* проедет λ_{ij} пассажиров, которое может принимать любое целочисленное значение на отрезке:

$$0 \le \lambda_{ij} \le min \lfloor a_{ij}, b_j \rfloor. \tag{5}$$

Такая постановка обоснована тем, что при решении задачи (3–4) величина поездок пассажиров (x_{ij}) между конкретными ОП *i* и *j* на маршруте может принимать целые неотрицательные значения только на отрезке (5). Любое число вне этого отрезка не может быть принято в качестве решения задачи (3–4).

Решение данной вероятностной задачи заключается в следующем. Обозначим через *A* событие, что от i-го до j-го ОП проехало λ_{ij} пассажиров, которое равно любому целому числу на отрезке (5). Тогда по классической формуле определения вероятности имеем, что

$$P(A) = \frac{N_A}{N}, \tag{6}$$

где *N* – количество равновозможных и несовместных групп в различных комбинациях, которые могут образовываться при выходе из салона ТС, число таких групп равно числу сочетаний из Q_{j-1} по b_j, т.е. $N = C_{Q_{j-1}}^{b_j}$;

N_A – количество равновозможных и несовместных групп в различных комбинациях, которые могут образовываться при выходе из салона ТС и соответствовать событию *A*, число таких групп равно числу сочетаний из a_{ij} по λ_{ij} ($C_{a_{ij}}^{\lambda_{ij}}$), а каждому такому сочетанию соответствуют сочетания из остальных

$(Q_{j-1}-a_{ij})$ по ($b_j-\lambda_{ij}$) в различных комбинациях ($C_{Q_{j-1}-a_{ij}}^{b_j-\lambda_{ij}}$). И тогда общее число благоприятствующих случаев событию *A* будет равно $N_A = C_{a_{ij}}^{\lambda_{ij}} C_{Q_{j-1}-a_{ij}}^{b_j-\lambda_{ij}}$.

Окончательно получим, что искомая вероятность события *A* определяется по формуле:

$$P(A) = P_{b_j}(\lambda_{ij}) = \frac{C_{a_{ij}}^{\lambda_{ij}} C_{Q_{j-1}-a_{ij}}^{b_j-\lambda_{ij}}}{C_{Q_{j-1}}^{b_j}}. \qquad (7)$$

По сути своей величина λ_{ij} является дискретной случайной величиной, которая принимает целые неотрицательные значения на отрезке (5). И так как для каждого её значения можно вычислить вероятность по формуле (7), то это соответствие можно представить в виде таблицы 2, которое указывает на то, что случайная величина λ_{ij} подчиняется данному закону распределения [3].

Таблица 2

Закон распределения дискретной случайной величины

Значения λ_{ij}	0	1	2	…	k
Вероятности p_l	p_0	p_1	p_2	…	p_k

Причем $k = min\ \lfloor a_{ij}, b_j \rfloor$, а вероятности p_l, где $0 \leqslant l \leqslant k$, определяются по формуле (6) для конкретного значения λ_{ij}.

В этом случае за решение задачи (3–4) из всех целых чисел λ_{ij} на отрезке (5) в качестве единственной искомой величины будем принимать то значение, для которой вероятность достигает своего максимального значения по аргументу λ_{ij}:

$$x_{ij} = maxP_{b_j}(\lambda_{ij})$$

Раз x_{ij} является наивероятнейшим числом среди всех значений случайной величины λ_{ij}, то для двух рядом стоящих чисел $(x_{ij}-1)$ и ($x_{ij}+1$) должны выполняться следующие неравенства [3]:

$$\frac{P_{b_j}(x_{ij}-1)}{P_{b_j}(x_{ij})} \le 1 \text{ и } \frac{P_{b_j}(x_{ij})}{P_{b_j}(x_{ij}+1)} \ge 1.$$

Раскрывая эти неравенства, приходим к следующим выражениям:

$$\frac{P_{b_j}(x_{ij}-1)}{P_{b_j}(x_{ij})} = \frac{x_{ij}(Q_{j-1}-a_{ij}-b_j+x_{ij})}{(a_{ij}-x_{ij}+1)(b_j-x_{ij}+1)} \leqslant 1$$

$$\frac{P_{b_j}(x_{ij})}{P_{b_j}(x_{ij}+1)} = \frac{(x_{ij}+1)(Q_{j-1}-a_{ij}-b_j+x_{ij}+1)}{(a_{ij}-x_{ij})(b_j-x_{ij})} \geqslant 1.$$

Решение же этих неравенств относительно x_{ij} приводит к следующему результату:

$$\frac{(a_{ij}+1)(b_j+1)}{Q_{j-1}+2}-1 \leqslant x_{ij} \leqslant \frac{(a_{ij}+1)(b_j+1)}{Q_{j-1}+2}.$$

А при больших значениях a_{ij} , b_j и Q_{j-1} данное выражение можно записать в следующем виде:

$$\frac{a_{ij}b_j}{Q_{j-1}}-1 \le x_{ij} \le \frac{a_{ij}b_j}{Q_{j-1}}.$$

Из этого двойного неравенства следует, что для вычисления элементов x_{ij} можно применять следующую формулу, округляя значения до целого числа:

$$x_{ij} = \frac{a_{ij}b_j}{Q_{j-1}}.$$

Заключение. Точность вычисления значений x_{ij} по данной формуле осуществлялась следующим образом. Используя таблицы матриц маршрутных корреспонденций пассажиропотоков (МКП) – $Y = (y_{ij})$, полученные в результате обработки материалов обследования пассажиропотоков во Владивостоке (2009 г.), брали данные входа и выхода, вычисляли элементы x_{ij} и на их основе формировали расчётные таблицы матриц МКП – $X = (x_{ij})$. Затем элементы этих матриц сравнивались между собой следующим способом.

Вводилась величина Θ – допустимая ошибки расчётов, которая зависит от частоты движения ТС, погодных условий и от времени в пределах суток (в часы пик величина допустимой ошибки будет выше).

Подсчитывалось количество элементов МКП матрицы X (допустим, M), для которых выполнялись следующие условия:

$$|x_{ij}-y_{ij}| > \theta \text{ и } \frac{\min(x_{ij}, y_{ij})}{\max(x_{ij}, y_{ij})} < \frac{\theta-1}{\theta}.$$

Зная, что число элементов любой матрицы МКП равно $n(n+1)/2$, то величина ошибки (W) определялась по формуле:

$$W = \frac{2M}{n(n+1)}100,$$

где n – число ОП на маршруте.

Для экспериментов брались матрицы как с рейсовыми, так и часовыми данными входа и выхода. Результаты сравнения представлены в табл. 3 (для рейсовых данных при $\Theta=3$ и для часовых данных при $\Theta=7$), а графически это отображено на рис. 1.

Таблица 3

Распределение оценок сравнения расчётного способа определения элементов маршрутных корреспонденций пассажиропотоков по интервалам

Интервал	Частота	
	абсолютная	относительная, %
0–1	12	5,7
1–2	15	7,1
2–3	25	11,8
3–4	24	11,4
4–5	12	5,7
5–6	15	7,1
6–7	21	9,9
7–8	26	12,3
8–9	35	16,6
9–10	13	6,2
10–11	5	2,3
11–12	4	1,9
12–13	0	0,0
13–14	1	0,6
свыше 14	3	1,4
Итого	211	100,0

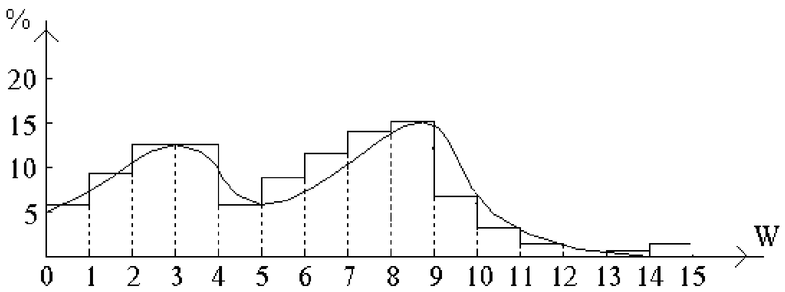

Рис. 1. Распределение частот

В связи с тем, что в результате кривая распределения на графике оказалась двухвершинной (или бимодальной), то мода непригодна для характеристики среднего положения оценки сравнения.

Тогда из всей совокупности матриц пришлось выделить отдельно по рейсам и отдельно по часам: по рейсам результаты сравнения отражены в табл. 4 и графически соответственно на рис. 2, а по часам – соответственно в табл. 5 и графически на рис. 3.

Таблица 4

Распределение оценок сравнения расчётного способа определения рейсовых элементов маршрутных корреспонденций пассажиропотоков по интервалам

Интервал	Частота	
	абсолютная	относительная, %
0–1	5	8,8
1–2	9	15,8
2–3	16	28,1
3–4	14	24,6
4–5	4	7,0
5–6	2	3,5
6–7	2	3,5
7–8	1	1,7
8–9	2	3,5
9–10	0	0,0
свыше 10	2	3,5
Итого	57	100,0

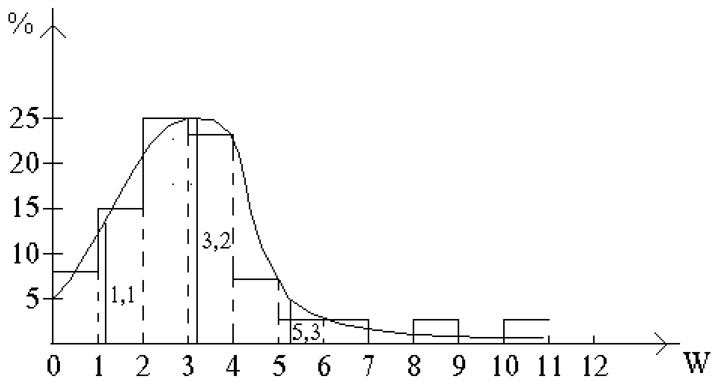

Рис. 2. Распределение частот

Таблица 5

Распределение оценок сравнения расчётного способа определения часовых элементов маршрутных корреспонденций пассажиропотоков по интервалам

Интервал	Частота	
	абсолютная	относительная, %
0–1	7	4,5
1–2	6	3,9
2–3	9	5,8
3–4	10	6,5
4–5	8	5,2
5–6	13	8,4
6–7	19	12,3
7–8	25	16,3
8–9	34	22,2
9–10	11	7,1
10–11	5	3,2
11–12	3	1,9
свыше 12	4	2,7
Итого	154	100,0

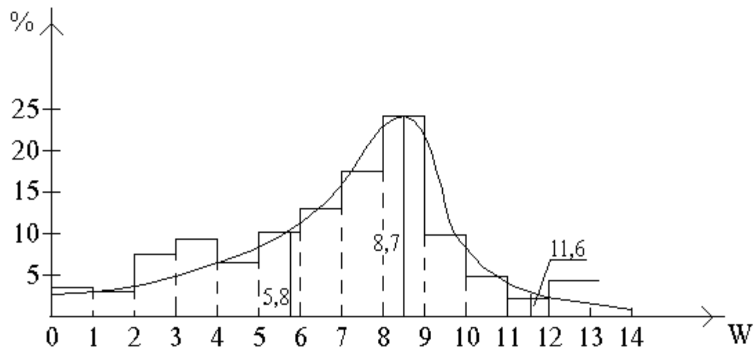

Рис. 3. Распределение частот

Анализ результатов сравнения показал, что отклонение численных значений колеблется в среднем в пределах 3,2% для рейсовых и 8,7% для часовых элементов МКП. Это указывает на то, что точность вычисления элементов матрицы

МКП выше при небольших значениях данных входа (a_i , $1 \le i \le n$) и выхода (b_j , $1 \le j \le n$), характеризующих слабые пассажиропотоки.

Важность этого вывода необходимо предусмотреть при разработке системы механизации получения данных входа-выхода и дальнейшей их автоматизации обработки. В этом случае комплекс математических моделей и алгоритмов обработки данных входа-выхода должен работать в такой последовательности: вначале вычислять элементы рейсовых матриц МКП, а затем на их основе определять параметры пассажиропотоков, необходимые для решения задач в управлении перевозками пассажиров в городах, в том числе матрицы МКП по каждому часу и за сутки в целом.

Следовательно, в настоящее время имеется возможность механизировать процессы сбора и обработки данных о входе-выходе пассажиров на каждом ОП маршрута, а затем автоматизировать их обработку с получением расчётным способом элементов МКП.

1. Андреев К.П., Терентьев В.В., Кулик С.Н. Обследование пассажиропотоков на городских автобусных маршрутах // Новая наука: проблемы и перспективы. 2016. № 2. С. 159–161.
2. Андреев К.П., Терентьев В.В. Пассажирские перевозки и оптимизация городской маршрутной сети // Мир транспорта. 2017. № 2. С. 159-161.
3. Гмурман В.Е. Теория вероятностей и математическая статистика. М.: Юрайт, 2019. 479 с.
4. Ембулаев В.Н., Дегтярёва О.Г., Белозерцева Н.П. Системный подход в теории и практике организации городских пассажирских перевозок. Владивосток: Изд-во ВГУЭС, 2013. 220 с.
5. Носов А.Л. Управление качеством работы городского пассажирского транспорта с использованием транспортной модели // Логистика сегодня. 2015. № 1. С. 38–47.
6. Фасхиев Х.А. Эффективный автобус для городских перевозок: как осуществлять выбор? // Логистика сегодня. 2017. №4. С. 274–288.
7. Федоров В.А. Современные задачи и проблемы натурных обследований пассажиропотоков (на примере Санкт-Петербурга) // Молодой учёный. 2015. №2. С. 333–342.

Транслитерация

1. Andreev K.P., Terent'ev V.V., Kulik S.N. Obsledovanie passazhiropotokov na gorodskih avtobusnyh marshrutah // Novaya nauka: problemy i perspektivy. 2016. № 2. P. 159–161.
2. Andreev K.P., Terent'ev V.V. Passazhirskie perevozki i optimizaciya gorodskoj marshrutnoj seti // Mir transporta. 2017. № 2. P. 159–161.
3. Gmurman V.E. Teoriya veroyatnostej i matematicheskaya statistika. M.: YUrajt, 2019. – 479 p.
4. Embulaev V.N., Degtyaryova O.G., Belozerceva N.P. Sistemnyj podhod v teorii i praktike organizacii gorodskih passazhirskih perevozok. Vladivostok: VGUES, 2013. 220 p.
5. Nosov A.L. Upravlenie kachestvom raboty gorodskogo passazhirskogo transporta s ispol'zovaniem transportnoj modeli // Logistika segodnya. 2015. №1. P. 38–47.
6. Faskhiev H.A. Effektivnyj avtobus dlya gorodskih perevozok: kak osushchestvlyat' vybor? // Logistika segodnya. 2017. №4. P. 274–288.

7. Fedorov V. A. Sovremennye zadachi i problemy naturnyh obsledovanij passazhiropotokov (na primere Sankt-Peterburga) // Molodoj uchyonyj. 2015. №2. P. 333–342.

Для цитирования: Ембулаев В.Н. Вероятностный метод определения поездок пассажиров на маршруте по данным входа и выхода // Территория новых возможностей. Вестник Владивостокского государственного университета экономики и сервиса. 2019. Т. 11, № 2. С. 58–69.

For citation: Embulaev V.N. A probabilistic method of determining a travel route according to the input and output, *The Territory of New Opportunities. The Herald of Vladivostok State University of Economics and Service*, 2019, Vol. 11, № 2, pp. 58–69.

DOI dx.doi.org/10.24866/VVSU/2073-3984/2019-2/058-069

Дата поступления: 04.04.2019.