

ТРАНСПОРТ

УДК 656.072.4.004.942

DOI <https://doi.org/10.32838/2663-5941/2020.4/32>

Жук М.М.

Національний університет «Львівська політехніка»

Півторак Г.В.

Національний університет «Львівська політехніка»

Гіць І.І.

Національний університет «Львівська політехніка»

Козак М.М.

Національний університет «Львівська політехніка»

ПРОГНОЗУВАННЯ ВИБОРУ ВИДУ ТРАНСПОРТУ У РАЗІ МІСЬКИХ ПЕРЕМІЩЕНЬ НА ОСНОВІ КЛАСИФІКАЦІЙНИХ ДЕРЕВ РІШЕНЬ

Оцінка вибору виду транспорту у разі переміщень є одним з важливих етапів транспортного планування. Під час міських переміщень основний вибір здійснюється між громадським (ГТ) та приватним транспортом (ПТ). Для прогнозування такого вибору застосовуються різні види моделей та способи, одним з яких є побудова дерев рішень. У разі прогнозування розподілу переміщень за видами потрібно враховувати мету поїздки, оскільки вибір користувачів відрізняється залежно від мети переміщення, що підтверджується проведеними дослідженнями. У роботі розглядаються чотири види поїздок: робочі, відпочинкові, з метою покупок та до/з транспортного вузла. Атрибутами користувачів, які враховувалися, є вік, стать, вид зайнятості (працюючі та працюючі студенти для робочих поїздок і працюючі та непрацюючі – для всіх інших) та наявність власного автомобіля. Поїздки характеризувалися за тривалістю, частотою та днем виконання (робочий чи вихідний). Дані для аналізу отримано шляхом проведення інтернет-опитувань у м. Львові. На їх основі побудовано класифікаційні дерева рішень для вибору користувачами виду транспорту (між громадським та приватним) залежно від мети поїздки. Виявлено, що одні і ті ж чинники по-різному впливають на вибір користувача за різної мети переміщення. Якщо наявність власного авто очікувано є найважливішим чинником впливу для вибору між ГТ та ПТ для всіх поїздок, то далі чинники відрізняються. Наприклад, для пересувань з метою відпочинку частота поїздки має найменший вплив на здійснення вибору, а от для поїздок з/до транспортного вузла це наступний після наявності автомобіля чинник впливу. З допомогою отриманих дерев рішень можна на основі даних щодо структури населення та переважаючих точок притягання в районах міста прогнозувати частку користувачів, які добиратимуться певним видом транспорту до місць праці, відпочинку, торгових закладів чи транспортних вузлів.

Ключові слова: міські переміщення, мета переміщення, дерева рішень, вибір виду транспорту, прогнозування.

Постановка проблеми. Вибір населенням способу переміщення в межах міської території суттєво впливає на рівень завантаження міської транспортної мережі, на міську екологічну систему, на сприйняття міста як комфортного для його мешканців та гостей середовища.

Основними способами переміщення в містах є приватний та громадський транспорт. Згідно з дослідженнями міської мобільності, прове-

деними у Львові, 52% містян для переміщень вибирають громадський транспорт, ще 23% – автомобіль, 18% переміщуються пішки і решта припадає на велопереміщення, поїздки у таксі тощо [1]. Розподіл пасажирів між різними видами транспорту відбувається залежно від їх особистих характеристик, вартості переміщення, часу в дорозі, доступності для користувача різних варіантів руху тощо.

Спосіб переміщення навіть для конкретно вибраної особи не є сталою величиною і може змінюватися залежно від зміни різних умов. Одним з важливих чинників впливу тут є мета поїздки.

Аналіз останніх досліджень та публікацій. У роботі [2] автори використовують для оцінки вибору режиму переміщення випадкові моделі мінімізації смутку (random regret minimization models), які були введені в транспортні дослідження як альтернатива випадковим корисним моделям. Автори провели аналіз вибору типу переміщення на основі мети поїздки. Виявилось, що коли для поїздки важливим чинником є пунктуальність (наприклад, робоча поїздка), то користувачі негативно сприймають навіть невелику різницю в характеристиках різних типів переміщень. А наприклад, у поїздках з метою відпочинку толерантність до різниці в характеристиках є значно вищою. Для перевірки того, чи визначення порогових значень рівня байдужості покращує прогностичну здатність моделі вибору виду транспорту, проведено розрахунки на основі даних опитувань респондентів (Нідерланди). Автори в роботі розглядали три варіанти режиму руху: автомобіль, велосипед та рух пішки. Основою вибору, окрім мети поїздки (робота, навчання, відпочинок, покупки та подорожі), є час переміщення.

У роботі [3] для оцінки вибору режиму руху пасажирями, що очікують на автовокзалі (між громадським та приватним (таксі чи прокат) транспортом), авторами пропонується застосування моделі SEM-Nested Logit, яка використовується для пошуку причинно-наслідкового зв'язку між спостережуваними та прихованими змінними (в авторській моделі це соціально-демографічні характеристики користувача (вік, стать, щомісячний дохід, рівень освіти, наявність власного транспортного засобу), характеристики подорожі (мета поїздки, час у дорозі, частота руху автобуса та показники комфорту). Автори визначили граничні значення показників комфортності, за яких імовірність переходу з приватного на громадський транспорт сягає найбільшого значення.

Недоліками дискретних моделей вибору режиму руху є те, що вони не відображають динамічного процесу поведінки під час подорожей. Адаже на вибір можуть впливати такі чинники, як, наприклад, інформація, отримана вже під час подорожі. Цей недолік може компенсувати теорія еволюційних ігор (Evolutionary Game Theory), яка припускає, що мандрівники коригують свою поведінку під час подорожі, порівнюючи корисність своєї поточної поведінки із середнім значенням корисності всієї системи.

Моделі, основані на цій теорії, використовуються, наприклад, для оцінки вибору режиму переміщення у великих містах під час ранкових пікових періодів [4] чи ймовірності вибору громадського транспорту власниками автомобілів під тиском міської стратегії розвитку громадського транспорту, з одного боку, та зростання заторів – з іншого [5].

Інші дослідники (наприклад, [6]) розширили теорію випадкової корисності на основі розрахунку «порогу байдужості» – коли величина різниці корисності двох варіантів переміщення менша певного значення, користувачі не в змозі її побачити, і тоді вибір здійснюється випадковим чином.

Автори [7] об'єднали ці дві теорії і в своєму дослідженні поєднали концепцію порогу байдужості з теорією еволюційних ігор.

Формулювання цілей статті. Прогнозування на транспорті є обов'язковим елементом транспортного планування. Достовірність отриманих у результаті прогнозування даних впливає на правильність прийняття рішень щодо проектування, функціонування та подальшого розвитку транспортної системи.

Основна ціль дослідження – визначити основні показники, які впливають на вибір користувачами міської транспортної системи виду свого переміщення, та на основі проведених досліджень побудувати дерева рішень вибору виду транспорту для майбутнього прогнозування частки переміщень цими видами транспорту.

Виклад основного матеріалу. Одним з методів аналізу даних та подальшої побудови прогнозних моделей є дерева прийняття рішень, з допомогою яких можна встановити залежність цільової змінної від вхідних змінних.

Дерево рішень складається з «гілок» та «листоків». Листки, які зазвичай зображуються у формі прямокутників, містять атрибути, а гілки – це розгалуження, які відображають альтернативи. Кожна гілка – це можливий результат.

Для побудови оптимального дерева прийняття рішення потрібно на кожному кроці вибирати той показник, який найкраще характеризує цільову функцію. Критерієм такого вибору може бути показник ентропії множини A [8, ст. 43]:

$$H(A, S) = - \sum_{i=1}^s \frac{m_i}{n} \log \frac{m_i}{n}, \quad (1)$$

де S – певна властивість множини, відносно якої розраховується ентропія;

n – кількість елементів у множині A ;

m – кількість елементів, які володіють властивістю S ;

s – кількість можливих значень властивості S .

На основі розрахованих значень ентропії розраховується показник приросту інформації [8, ст. 45]:

$$Gain(A, Q) = H(A, S) - \sum_{i=1}^q \frac{|A_i|}{|A|} H(A_i, S), \quad (2)$$

де A_i – множина елементів A , в якій атрибут Q , який може мати q можливих значень, має значення i .

Щоб уникнути неточності результату через те, що атрибути мають неоднакову кількість значень, розраховується величина поправки показника приросту інформації [8, ст. 49]:

$$SplitInfo(A, Q) = - \sum_{i=1}^q \frac{|A_i|}{|A|} \log_2 \frac{|A_i|}{|A|} \quad (3)$$

$$GainRatio(A, Q) = \frac{Gain(A, Q)}{SplitInfo(A, Q)} \quad (4)$$

На кожному кроці побудови дерева рішень потрібно вибирати той атрибут, для якого скоригований приріст інформації найбільший.

На основі проведених опитувань користувачів транспортної системи Львова отримано статистичну інформацію щодо вибору ними способу свого пересування у разі виконання перемішень. Результати досліджень показали, що ймовірність вибору громадського транспорту залежить від мети поїздки: для робочих поїздки та поїздки з/до транспортного вузла громадський транспорт вибрали 71% користувачів, для відпочинкових – 48%, а для поїздки з метою покупок – 53%. Параметрами, які враховувалися, є соціально-демо-

графічні характеристики користувача (вік, стать, зайнятість та наявність власного автомобіля) та характеристики поїздки (тривалість поїздки і час або частота її виконання).

Щодо загальної структури вибірки, то вона складається з 48% чоловіків та 52% жінок. 47% опитаних молодші 21 року, решта – старші (мінімальний вік опитаних – 17 років, максимальний – 56 років), 26% мають власний автомобіль. Працюючих – 49%, ще 12% – одночасно і працюючі, і студенти.

Щодо характеристик поїздки, то вони відрізняються залежно від мети поїздки. За частотою поїздки було поділено на дві групи: ті, які відбуваються кілька разів на місяць, і ті, які відбуваються раз на кілька місяців. Кілька разів на місяць здійснюються 58% відпочинкових поїздки та 42% поїздки з/до транспортного вузла. Щодо дня тижня, то в робочі дні (понеділок–п’ятниця) відбувається 33% поїздки з метою покупок, 48% поїздки до/з транспортного вузла та 23% відпочинкових поїздки. Розподіл поїздки за тривалістю подано на рис. 1.

Початкова ентропія для ситуації вибору користувачем громадського транспорту становитиме:

- під час робочих поїздки: $H(A, GT) = 0,875$;
- під час відпочинкових поїздки: $H(A, GT) = 0,998$;
- під час поїздки з метою покупок: $H(A, GT) = 0,998$;
- під час поїздки з/до транспортного вузла: $H(A, GT) = 0,869$.

Результати розрахунків приростів інформації з подальшим коригуванням для всіх атрибутів, які розглядаються, подано в таблиці 1.

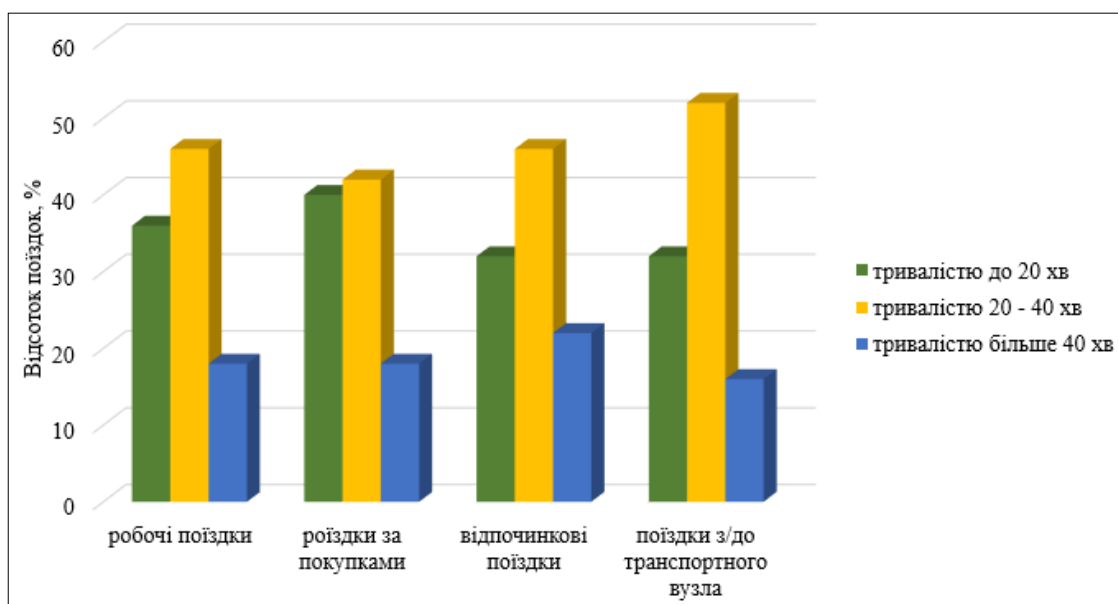


Рис. 1. Розподіл поїздки за тривалістю залежно від мети переміщення

**Розрахунок приростів інформації для атрибутів у разі прогнозування вибору
для поїздки громадського транспорту**

Атрибут Q	$Gain(A, Q)$	$SplitInfo(A, Q)$	$GainRatio(A, Q)$	Атрибути впливу, впорядковані в порядку спадання їх значимості
Робочі поїздки				
Вік	0,110	1,0	0,110	Наявність авто
Стать	0,055	0,982	0,056	Вік
Зайнятість	0,111	0,999	0,110	Зайнятість
Наявність авто	0,703	1,204	0,584	Стать
Тривалість поїздки	0,014	1,490	0,009	Тривалість поїздки
Відпочинкові поїздки				
Вік	0,085	0,998	0,085	Наявність авто
Стать	0,018	0,998	0,018	Зайнятість
Зайнятість	0,121	0,983	0,124	Вік
Наявність авто	0,601	0,998	0,602	Тривалість поїздки
Тривалість поїздки	0,051	1,523	0,034	Стать
Частота поїздки	0,012	0,983	0,012	Частота поїздки
Поїздки з метою покупок				
Вік	0,038	0,999	0,038	Зайнятість
Стать	0,028	0,998	0,028	Вік
Зайнятість	0,063	0,973	0,065	Тривалість поїздки
День тижня	0,016	0,918	0,017	Стать
Тривалість поїздки	0,050	1,494	0,033	День тижня
Поїздки до/з транспортного вузла				
Вік	0,095	1,0	0,095	Наявність авто
Стать	0,0003	0,999	0,0003	Частота поїздки
Зайнятість	0,022	0,981	0,023	Вік
Наявність авто	0,321	0,823	0,390	Зайнятість
Тривалість поїздки	0,021	1,443	0,014	Тривалість поїздки
Частота поїздки	0,096	0,984	0,097	День тижня
День тижня	0,002	0,999	0,002	Стать

На основі отриманих даних будуються дерева рішень для вибору виду транспорту – громадського або приватного (рис. 2).

Знаючи характеристики користувача, з допомогою отриманих дерев рішень можна спрогнозувати, який вид транспорту він вибере у разі поїздки з тією чи іншою метою.

Висновки. Основні висновки проведеного дослідження.

1. Визначено, що залежно від мети поїздки атрибути мають різний вплив на вибір виду транспорту у разі міських переміщень: зокрема, для робочих поїздок найбільший вплив на здійснення вибору має наявність авто та вік користувача, а найменший – тривалість поїздки; для відпочинкових найважливішими атрибутами є наявність авто та зайнятість, а найменш важливим – частота

поїздки; для поїздок з метою покупок найбільший вплив на вибір виду транспорту має зайнятість та вік, а найменший – день тижня; і для поїздок з/до транспортного вузла найзначимішими атрибутами є наявність авто та частота поїздки, а найменш значимим – стать користувача.

2. Для прогнозування вибору виду транспорту пропонується метод побудови дерева рішень, який є відносно простим у використанні та досить ефективним [9, ст. 442].

3. З допомогою отриманих дерев рішень можна на основі даних щодо структури населення та переважаючих точок притягання в районах міста прогнозувати частку користувачів, які добиратимуться певним видом транспорту до місць праці, відпочинку, торгових закладів чи транспортних вузлів.

1 – вибір користувачем громадського транспорту
0 – вибір користувачем приватного автотранспорту

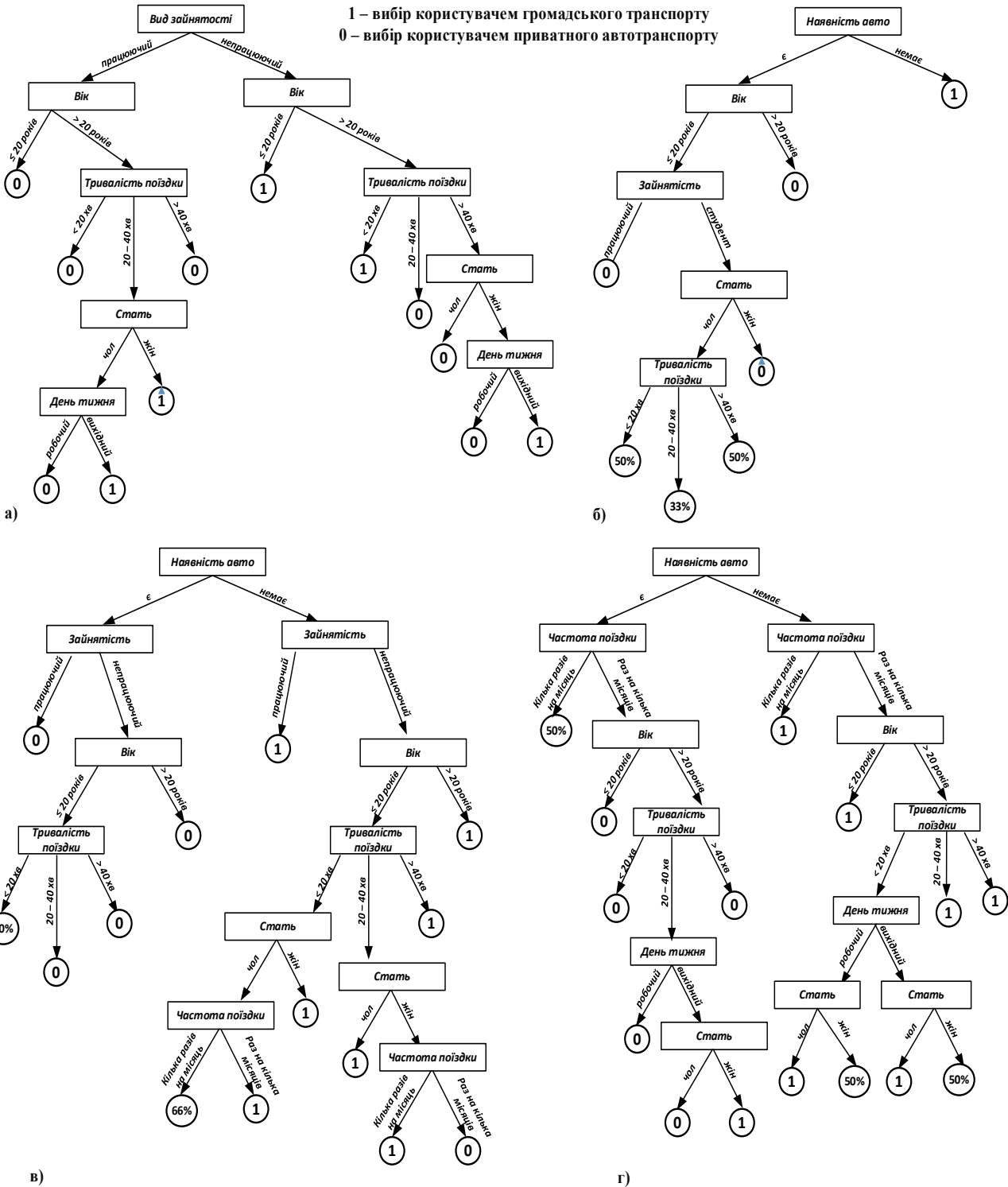


Рис. 2. Древа рішень для прогнозування вибору користувачем транспортних послуг виду транспорту: а) пересування з метою покупок; б) пересування з робочою метою; в) пересування з метою відпочинку; г) пересування до/з транспортного вузла

Список літератури:

1. Веб-сторінка «Мобільність Львова». URL: <https://mobilitylviv.com/>.
2. Sunghoon Jang, Soora Rasouli, Harry Timmermans. Tolerance and Indifference Bands in Regret–Rejoice Choice Models: Extension to Market Segmentation in the Context of Mode Choice Behavior. *Transportation Research Record: Journal of the Transportation Research Board*. 2018. Volume 2672, Issue 47. P. 23–34. URL: <https://journals.sagepub.com/doi/full/10.1177/0361198118787629>.
3. Yan Han, Wanying Li, Shanshan Wei and Tiantian Zhang. Research on Passenger’s Travel Mode Choice Behavior Waiting at Bus Station Based on SEM-Logit Integration Model. *Sustainability*. 2018. Volume 10(6), 1996. URL: <https://www.mdpi.com/2071-1050/10/6/1996/htm#B45-sustainability-10-01996>.
4. Chaoqun Wu, Yulong Pei, and Jingpeng Gao. Evolution Game Model of Travel Mode Choice in Metropolitan. *Discrete Dynamics in Nature and Society*. 2015. Volume 2015, Article ID 638972, 11 pages. URL: <https://www.hindawi.com/journals/ddns/2015/638972/>.
5. Huawei Gong and Wenzhou Jin. Analysis of Urban Car Owners Commute Mode Choice Based on Evolutionary Game Model. *Journal of Control Science and Engineering*. 2015. Volume 2015, Article ID 291363, 5 pages. URL: <https://www.hindawi.com/journals/jcse/2015/291363/>.
6. Wang W., Wang D.Z.W., Sun H.J. and Wu J.J. Public transit service operation strategy under indifference thresholds-based bi-modal equilibrium. *Journal of Advanced Transportation*. 2016. Volume 50, No. 6. Pp. 1124–1138. URL: <https://doi.org/10.1002/atr.1393><https://onlinelibrary.wiley.com/doi/full/10.1002/atr.1393>.
7. Xinjie Zhang, Hongzhi Gao. Research on Travel Mode Choice Behaviors Based on Evolutionary Game Model Considering the Indifference Threshold. *IEEE Access*. 2019. Volume 7. Pp. 174083–174091. URL: <https://ieeexplore.ieee.org/abstract/document/8918435/references#references>.
8. Марченко О.О., Россада Т.В. Актуальні проблеми Data Mining : навчальний посібник для студентів факультету комп’ютерних наук та кібернетики. Київ, 2017. 150 с.
9. Ситник В.Ф., Ситник Н.В. Деревя рішень у системах дейтамайнінгу. *Формування ринкової економіки* : зб. наук. пр. Київ. нац. екон. ун-т ім. В. Гетьмана. Київ, 2006. Вип. 16. С. 442–454.

Zhuk M.M., Pivtorak H.V., Gits I.I., Kozak M.M. FORECASTING THE MODE CHOICE OF TRANSPORT FOR URBAN MOVEMENTS USING CLASSIFICATION DECISION TREES

Evaluation of the mode choice of transport when moving is one of the important stages of transport planning. For urban transport, the main means of transport is public (PuT) and private transport (PrT). Different types of models and methods are used to forecasting this choice, one of which is the construction of decision trees. When forecasting the distribution of movements by type, you need to consider the trip purpose, because the choice of users differs depending on the movement’s purpose. This is confirmed by research. The paper considers four types of trips: work, leisure, shopping and to/from the transport hub. The following user attributes were taken into account: age, gender, type of employment (working and working students for work trips and working and non-working – for everyone else) and the availability of their own car. Trips are characterized by duration, frequency and day of performance (working or weekend). Data for analysis were obtained by conducting online surveys in Lviv city. Based on them, classification decision trees are built for users to choose the type of transport (between public and private) depending on the purpose of the trip. It was found that the same factors affect the user’s choice in different ways of transportation. Having your own car is expected to be the most important influencing factor for choosing between PuT and PrT for all trips. The influence of other factors is different. For example, for leisure travel, the frequency of travel has the least influence on the choice, but for travel to/from the transport hub is the next impact factor after the presence of a car. With the help of the obtained decision trees, it is possible to predict the share of users who will choose by a certain type of transport to places of work, rest, shopping or transport hubs on the basis of data on population structure and prevailing points of attraction in the city districts. With the help of the obtained decision trees, it is possible to predict the share of users who will choose by a certain type of transport to places of work, rest, shopping center or transport hubs on the basis of data on population structure and prevailing points of attraction in the city zones.

Key words: urban movements, trip purpose, decision trees, mode choice, forecasting.