

КЛАСТЕРНИЙ АНАЛІЗ ЗУПИННИХ ПУНКТІВ ТРАНСПОРТНО-ПЕРЕСАДОЧНИХ ВУЗЛІВ МГПТ

В.О. Вдовиченко

*Харківській національній автомобільно-дорожній університет,**Vval2301@gmail.com*

В умовах обмеженої класифікації зупинних пунктів (ЗП) транспортно-пересадочних вузлів (ТПВ) та відсутності їх типового розподілу за показниками формування маршрутного потоку доцільним є проведення їх кластерного аналізу. Кластерний аналіз є методом класифікаційного розподілу який в межах виділеного предмету дослідження дозволяє розділити базові ЗП ТПВ на відповідні однорідні групи за класифікаційними ознаками відносно параметрів вхідних маршрутних потоків [1]. Кластерний аналіз ЗП ТПВ передбачає реалізацію наступних етапів: формування сукупності номерів (збір вихідної інформації про фактичні параметри вхідних маршрутних потоків у ЗП), виділення сукупності об'єктів (розподіл за групами чинників), формування множини алгоритмів кластеризації, побудова графіків розподілу номерів на сукупності об'єктів, виділення кластерних груп ЗП та їх характеристичних ознак. Для проведення кластерного аналізу були використані натурні спостереження за параметрами формування маршрутних потоків в ЗП ТПВ міст України: Харкова, Кривого Рогу, Херсона.

Кластеризація передбачає проведення процедури оцінювання адаптованості розподілу до визначення діапазонів технологічних параметрів формування простою транспортних засобів (ТЗ) у ЗП. Для цього використаний метод k -середніх. Процедура розподілової кластеризації, заснована на розбитті безлічі елементів векторного простору на заздалегідь визначену кількість кластерів. Метод відноситься до неієрархічних алгоритмів кластеризації та являє собою ітераційну процедуру, що складається з наступних етапів: вибирається число кластерів, з вихідної безлічі даних випадковим чином обирають базові записи, які будуть служити початковими центрами кластерів, для кожного запису вихідної вибірки визначається найближчий до неї центр кластера, обчислюються центроїди – центри тяжкості кластерів, центр кластера зміщується в його центр ваги.

Для встановлення параметрів кластеризації використовувались середні значення часу простою ТЗ у черзі, часу маневрування, часу технологічного простою, часу сервісного простою, загальна кількість конфліктних ситуацій у ЗП. Проведення якісного розподілу параметрів за групами чинників впливу дозволяє значно знизити рівень невизначеності та скоротити розрахункову процедуру кластеризації. Для цього виділені групи факторних ознак, що за характером протікання процесів можуть впливати на параметрів маршрутного потоку та параметрів простою ТЗ. До таких ознак відноситься: кількість маршрутів (r_n), кількість рейсів (k_r^{sp}), пасажирообмін ТЗ (p_{en-ex}^{sp}), довжина ділянки маршруту до ТПВ (l_r^{sp}), довжина посадочного майданчика ЗП (l_{sp}). Об'єднання сукупності параметрів вхідного маршрутного потоку та груп чинників впливу дає можливість визначити загальну варіантність двомірних алгоритмів кластеризації. Алгоритм кластеризації представляє собою функцію $a_{ij}: O_j \rightarrow P_i$ яка кожній групі чинників (O_j) ставить у відповідність мітку кластера (P_i). Процедура ітерації розподілу за кластерами

припиняється, коли їх кордони не перестануть змінюватися від ітерації до ітерації. Критерієм оцінювання якості алгоритму кластеризації є показник ефективності розподілу [2]

$$I_e = \sum_{j=1}^c \sum_{i=1}^m u_{ji} \left(d(k_j, \overline{x_j}) - d(c_i, x_i) \right)^2, \quad (1)$$

де c – кількість кластерів;

m – загальна кількість вхідних номерів;

u_{ij} – елемент матриці приналежності номерів до кластерів;

k_j – центр номерів кластеру j ;

$\overline{x_j}$ – середнє значення множини міток кластеру j ;

c_i – значення номеру i ;

x_i – значення мітки i .

Структурна адаптація передбачає для кожного типу номерів прямий перебір факторної ознаки шляхом випадкового пошуку з нелінійною тактикою. Кожен новий варіант алгоритму кластеризації характеризує зміну індексу ефективності

$$I_e[a+1] = I_e[a] + \Delta I_e[a+1], \quad (2)$$

де $I_e[a]$ – індекс ефективності розподілу для варіанта a ;

$\Delta I_e[a+1]$ – прирощення індексу ефективності для варіанта $a+1$.

У разі, якщо наступний крок при переборі призводить до зниження значення індексу ефективності, то відбувається повернення до попереднього варіанта. Оператор прийняття кроку має бінарну форму

$$Z[a+1] = \begin{cases} 1, & \text{при } \Delta I_e[a+1] \leq 0 \\ 0, & \text{при } \Delta I_e[a+1] > 0 \end{cases}, \quad (3)$$

де 1 – повернення до варіанта a ; 0 – прийняття варіанта $a+1$.

Шляхом перебору можливих алгоритмів відділені варіанти, що забезпечують максимальне значення індексу кластеризації, а саме: $a_{13} : p_{en-ex}^{sp} \rightarrow t_r^{dt}$, $a_{25} : r_n \rightarrow N_{rl}^{sp}$, $a_{31} : k_r^{sp} \rightarrow t_r^{wq}$, $a_{42} : l_{sp} \rightarrow t_r^m$, $a_{54} : l_r^{sp} \rightarrow t_r^{ad}$.

Література.

1. Vdovychenko V. Cluster analysis of transport interchanges of urban public passenger transport. The development of technical sciences: problems and solutions: Conference Proceedings - Brno, European Network for Academic Integrity, 2018. p. 119-122.

2. Гитис Л.Х. Статистическая классификация и кластерный анализ. Москва: Издательство МГГУ, 2003. 157 с.