

УДК 656.022.5

ВИЗНАЧЕННЯ ДОБОВИХ ПЕРІОДІВ ЗМІНИ ЧАСУ ОБЕРТУ НА МАРШРУТІ З ДОПОМОГОЮ КЛАСТЕРНОГО АНАЛІЗУ

DEFINING THE DAY TIME PARTS OF ROUTE CYCLE TIME CHANGING WITH USING THE CLUSTER ANALYSIS

Галина Півторак, Андрій Білоус

Національний університет «Львівська політехніка»,
79013, м. Львів, вул. С. Бандери, 12

This thesis is dedicated to investigation of the modern and efficient algorithm of appropriate day time periods for transit route work designing. In the frame of apparent time period the route cycle time is considered stable and constant. As input data for the algorithm the GPS tracking data of route vehicles are used. The time of staying at the transit stop is the key factor for defining the time periods. For the main purpose of the algorithm the cluster analysis is used.

Однією з основних ознак якості роботи рухомого складу на міських маршрутах громадського транспорту є дотримання ним розкладу руху та вчасна корекція розкладів відповідно до вимог користувачів. Згідно з багатьма зарубіжними дослідженнями, ефективним інструментом, який допомагає контролювати дотримання розкладу руху, є метод вибору певних зупинок маршруту в якості контрольних точок, на яких відстежується дотримання розкладу руху. За отриманими результатами відбувається корекція розкладу руху на маршруті [1, 2].

Основними показниками, за якими визначають приналежність зупинки маршруту до контрольної точки, є час простою на зупинці та відхилення фактичного часу прибуття на зупинку відносно запланованого розкладом. Оскільки час простою на зупинці прямо залежить від кількості пасажирів, що здійснюють посадку/висадку, що визнано багатьма дослідженнями [3, 4], то коректним є робити висновок про пасажиронапруженість зупинки на основі часу простою. Проте числові значення пасажиронапруженості для однієї і тієї ж зупинки відрізнятимуться протягом доби. Тобто, для проведення подальшого аналізу доцільно поділити добу роботи маршруту на окремі часові періоди.

Поділ доби роботи маршруту на пікові та міжпікові періоди є загальноприйнятим. Проте конкретні часові рамки цих періодів на різних маршрутах можуть відрізнятися. Зручним та універсальним інструментом для визначення меж часових періодів є кластерний аналіз. В якості критерію схожості кластерів використовується час простою на зупинці (хх:сс) та конкретний час (гг:хх:сс) перебування автобуса на цій зупинці.

Для проведення кластеризації використовується метод нечітких C -середніх [5]. Метод полягає в мінімізації функції:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, 1 \leq m \leq \infty \quad (1)$$

де m - експонентна вага розрахунку цільової функції і центрів кластерів (як правило, $m = 2$); u_{ij} - ступінь належності x_i до кластеру j ; x_i - набір даних; c_j - центр кластера.

Нечітке розділення відбувається за допомогою ітеративної оптимізації нечіткої функції з одночасним оновленням ступеня належності та центру кластера. Ітерація зупиняється тоді, коли:

$$\max_{ij} \left\{ \left| u_{ij}^{(k+1)} - u_{ij}^{(k)} \right| \right\} < \varepsilon \quad (2)$$

де k - кількість ітерацій; ε - параметр збіжності ($\varepsilon = 0 \div 1$); x_i - набір даних.

Чим більша кількість операцій і ближчий до нуля параметр збіжності, тим кращим (точнішим) є отриманий результат [6].

Для проведення кластеризації необхідно попередньо задати бажану кількість кластерів, максимальну кількість ітерацій, параметр збіжності методу, а також експонентну вагу розрахунку цільової функції і центрів кластерів.

Набори даних для проведення кластерного аналізу отримані за допомогою GPS-даних транспортних засобів, які обслуговують маршрути громадського транспорту. Отримання даних забезпечується в онлайн режимі за допомогою API в JSON форматі. Зведені дані мають наступний вигляд (табл. 1):

Таблиця 1. Статистична інформація, отримана за допомогою GPS-даних ТЗ

День тижня	Дата	Час (гг:хх:сс)	Номер маршруту	Напрямок	Номер трекера	Географічна довгота	Географічна широта

Відгук API відбувається кожні 12 с, відповідно, отриманий масив статистичної інформації є занадто великим для опрацювання його без допомоги програмного середовища. Для реалізації алгоритму обчислень написано програму в програмному середовищі Python.

На початковому етапі весь масив даних поділено на окремі таблиці за окремими маршрутами. Потім виділено інформацію (рядки даних), яка стосується зупинок ГПТ. Час простою на зупинці визначається за формулою:

$$t_{np} = t_n - t_e \quad (3)$$

де t_n - час прибуття автобуса на зупинку; t_e - час відправлення автобуса з зупинки.

З подальшого розгляду також відкидалися кінцеві зупинки, оскільки час простою на них не залежить від пасажиропотоку, а регламентується розкладом руху автобуса.

Враховуючи дослідження коливань пасажирських потоків на маршруті [7], весь масив даних поділено на сім кластерів: передпіковий період, ранковий пік, перший міжпіковий період, обідній пік, другий міжпіковий період, вечірній пік та післяпіковий період. Після задання кластерів обирається центр кожного кластера. Центр можна обирати довільно, але правильний його вибір суттєво зменшує необхідну кількість ітерацій для досягнення прийняттого результату. Для проведення процедури за центри кластерів було обрано значення часу доби: 7:00 год (для першого кластера), 08:30 год (для другого кластера), 11:00 год (для третього кластера), 13:00 год (для четвертого кластера), 15:00 год (для п'ятого кластера), 18:00 год (для шостого кластера) та 20:30 год (для сьомого кластера). Подальший поділ відбувається за таким алгоритмом:

Нехай x_{ijk} - час простою k -ого автобуса на i -ій зупинці під час j -ого прибуття ($i = 1 \div n$, де n - кількість зупинок на маршруті; $j = 1 \div m$, де m - кількість прибуттів автобуса на i -ту зупинку; $k = 1 \div r$, де r - кількість автобусів, що працюють на даному маршруті). Якщо «відстань» між часом простою автобуса на зупинці x_{ijk} та центром кластера є мінімальною і,

водночас, година прибуття автобуса на зупинку однакова або відрізняється на одиницю, то дані об'єкти відносяться до одного кластера.

Подальша робота над алгоритмом використання контрольних точок маршруту проводитиметься в напрямку аналізу часу прибуття автобуса на зупинку та відхилення його від планового. Також доцільним є подальше дослідження меж кластерів засобами нечіткої логіки. Отримані результати дадуть можливість здійснювати трьохфакторний аналіз зупинок ГПТ (пасажирообмін, період доби та відхилення часу прибуття).

Література:

1. S.C. Wirasinghe, G. Liu, *Determination of the number and locations of time points in transit schedule design – Case f a single run. Annals of Operations Research, Canada, Volume 60, Issue 1, 1995, P.161 – 191. DOI 10.1007/BF02031945.*
2. James G. Strathman, Thomas J. Kimpel, Kenneth J. Dueker. *Bus Transit Operations Control: Review and an Experiment Involving Tri-Met's Automated Bus Dispatching System. Journal of Public Transportation. Volume 4, Issue 1, 2001, P. 1–26.*
3. Гудков В.А., Миротин Л.Б., Вельможин А.В., Ширяев С.А. / *Пассажирские автомобильные перевозки. Под ред. В.А. Гудкова. - М.: Горячая линия– Телеком, 2004. - 448 с.*
4. Доля В.К. *Теоретические основы и методы организации маршрутных автобусных перевозок пассажиров в крупнейших городах: Автореф. дис. д. т. н.: 05.22.10. - М.: МАДИ, 1993. - 42 с.*
5. Штовба С.Д. *Побудова функцій належності нечітких множин за кластеризацією експериментальних даних // Інформаційні технології та комп'ютерна інженерія. – 2006. – №2. – С. 92–95.*
6. Das S. *Pattern Recognition using the Fuzzy c-means Technique. International Journal of Energy, Information and Communications. Volume 4, Issue 1, 2013, P. 1–16.*
7. Давідіч Ю.О. *Розробка розкладу руху транспортних засобів при організації пасажирських перевезень: навч. посіб. / Ю.О. Давідіч; Харк. нац. акад. міськ. госп-ва. – Х.: ХНАМГ, 2010. – 345 с.*