

УДК 519.816

САМОРОДОВ В.Б., д.т.н., проф., НТУ «ХПІ»
АНДРОСЕНКО В.В., асс., НТУ «ХПІ»

ВИКОРИСТАННЯ КЛАСТЕРНИХ РАНЖИРУВАНЬ ДЛЯ ПРОГНОЗУВАННЯ ВИКОРИСТАННЯ РУХОМОГО СКЛАДУ

В роботі розроблено методику для виконання прогнозування використання рухомого складу на підприємствах автотранспортної галузі. Розроблено алгоритм, який включає послідовність процедур для проведення прогнозування та описано основні вхідні дані. Формалізовано та надано пояснень щодо всіх операцій прогнозування. Особливо представлено механізм для визначення сезонної складової динаміки використання рухомого складу.

Вступ. Транспортно-експедиційний ринок України набуває поступового розвитку, основною рисою якого є зв'язок між інформаційною та технологічною складовою транспортного процесу. Іноземний досвід свідчить про необхідність впровадження взаємопов'язаних інформаційно-технологічних систем для різних видів транспорту. Загальним підсумком такого зв'язку є чітка синхронізація перевізного процесу, що позитивно впливає як на економічний стан підприємств так і на якість обслуговування клієнтів. Однак, відсутність чіткого зв'язку між підприємствами в Україні призводить до ускладнення процедури синхронізації. В таких умовах перевізникам доводиться проводити самостійний багатофакторний пошук клієнтів з урахуванням наявного рухомого складу та можливості його додаткового залучення за рахунок приватних підприємств. Це призводить до необхідності проведення політики оптимізації власних транспортних потужностей під потреби великих клієнтів. Політика полягає в раціоналізації та уніфікації парку рухомого складу. Частіше його структура формується для нівелювання коливань попиту та постійного виконання перевезень протягом окремого терміну – сезону або року. Такий підхід викликає появу низки проблем, серед яких основне місце посідають питання прогнозування використання рухомого складу на наступний період.

Побудова універсальної методики прогнозування використання рухомого складу для автотранспортних підприємств дозволить автоматизувати процес прийняття оперативних рішень щодо формування оптимального парку рухомого складу. Результатом впровадження методики може стати система оперативної зміни структури парку за рахунок усвідомленої задачі або взяття транспортних засобів в оренду.

Актуальність проблеми дослідження. До найбільш розповсюджених моделей прогнозування використання рухомого складу відносяться статистичні моделі. В них, за допомогою аналізу часових рядів формуються закони розподілення випадкової величини використовуючи які виконують безпосередньо прогнозування. Недоліком використання статистично-ймовірнісних методів є необхідність визначення та обґрунтування фактичного закону розподілення випадкової величини [1, 2]. При цьому, частіше, вибір будь-якого з них вимагає згладжування окремих частот. Який вплив створює процедура відмови від реальної структури частот та нав'язування класичного закону залишається невідомим. Доволі складним стає і вибір критерію для перевірки розходження теоретичного та емпіричного розподілу

Чисельні методи, а саме методи інтерполяції та апроксимації [3], через значну простоту використання є більш прийнятними однак вони мають низку власних проблем

та недоліків. Серед них основне місце посідає вибір типу поліному. Неможливість його точного та чисельного обґрунтування призводить до появи “надії” стосовно правильно прийнятого рішення.

Побудова моделі на основі кластерних ранжирувань дозволить нівелювати ці недоліки та дасть можливість прийняти обґрунтоване рішення стосовно зміни структури парку у часі. Невідповідності в кластерних ранжируваннях нівелюються за допомогою аналізу ядра протиріч, що дозволяє спростити процедуру пошуку тренду та сезонної складової.

Рішення задачі. Прогнозування використання рухомого складу для транспортно-експедиційних підприємств внаслідок специфічності автотранспортних послуг стає причиною появи відповідних вимог серед яких виділено три основних:

- відстеження місячної динаміки;
- обов'язковість побудови загального тренду;
- необхідність побудови сезонного ранжирування.

У відповідності до представлених вимог розроблено алгоритм проведення прогнозування використання рухомого складу. Основою для наступного прогнозу є кількість годин, які було відпрацьовано відповідним автомобілем. В якості математичного апарату використовується метод погодження ранжирувань з кластерами [4]. Рациональну послідовність операцій наведено на рисунку 1.



Рисунок 1 - Послідовність операцій прогнозування використання рухомого складу

Математична постановка задачі виглядає наступним чином. Припустимо, що множина автомобілів, які включено до парку рухомого складу описується множиною $H = \{h_1, h_2, \dots, h_p\}$. Час виконання робіт за кожною маркою рухомого складу

описується бінарним відношенням (1). За вхідною інформацією формується множина ранжирувань з кластерами, які виглядають як послідовність $pos_1^{h_i} = (t_0 < t_1 = t_j < \dots)$.

$$TP^{h_i} = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & 12 \end{matrix} \\ \begin{matrix} t_1 \\ t_2 \\ \dots \\ t_j \end{matrix} & \begin{bmatrix} tp^{h_i}(t_1,1) & tp^{h_i}(t_1,2) & \dots & tp^{h_i}(t_1,12) \\ tp^{h_i}(t_2,1) & tp^{h_i}(t_2,2) & \dots & tp^{h_i}(t_2,12) \\ \dots & \dots & \dots & \dots \\ tp^{h_i}(t_j,1) & tp^{h_i}(t_j,2) & \dots & tp^{h_i}(t_j,12) \end{bmatrix} \end{matrix} \quad (1)$$

В результаті отримуємо множину ранжирувань з кластерами, які виглядають як послідовність $pos_1^{h_i} = (t_0 < t_1 = t_j < \dots)$. Всі ранжирування за окремим автомобілем та окремим місяцем зводяться до бінарного відношення за допомогою залежності

$$bo_{jm}^{h_i}(t_k, t_m) = \begin{cases} 1, & \text{якщо } t_k^{h_i} \leq t_m^{h_i} \\ 0, & \text{якщо } t_k^{h_i} > t_m^{h_i} \end{cases} \quad k \neq m, \quad k = 1, 2, \dots, j \quad m = 1, 2, \dots, j. \quad (2)$$

В результаті одержуємо матрицю

$$BO_{jm}^{h_i} = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & j \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \dots \\ j \end{matrix} & \begin{bmatrix} 1 & bo_{jm}^{h_i}(t_1, t_2) & \dots & bo_{jm}^{h_i}(t_1, t_j) \\ bo_{jm}^{h_i}(t_2, t_1) & 1 & \dots & bo_{jm}^{h_i}(t_2, t_j) \\ \dots & \dots & \dots & \dots \\ bo_{jm}^{h_i}(t_j, t_1) & bo_{jm}^{h_i}(t_j, t_2) & \dots & 1 \end{bmatrix} \end{matrix}, \quad (3)$$

де jm – місяць, для якого виконано ранжирування.

Наступний аналіз полягає у виділенні ядра протиріч. Для цього необхідно виконати попарне перемноження всіх елементів відповідних матриць. Ядром протиріч є пари для яких результат дорівнює нулю в обох напрямках, тобто

$$ads^{h_i}(t_k, t_m) = bo_{jm}^{h_i}(t_k, t_m) \cdot bo_{jm}^{h_i}(t_k, t_m) = bo_{jm}^{h_i}(t_m, t_k) \cdot bo_{jm}^{h_i}(t_m, t_k) = 0 \quad k \neq m \quad k = 1, 2, \dots, j \quad m = 1, 2, \dots, j \quad (4)$$

Підсумкове ядро протиріч визначається за допомогою математичної операції об'єднання окремих ядер протиріч ранжирувань з кластерами

$$ADS^{h_i} = \cup ads^{h_i}(t_k, t_m) \quad k \neq m \quad k = 1, 2, \dots, j \quad m = 1, 2, \dots, j. \quad (5)$$

В результаті одержуємо послідовність протиріч. Для їх усунення та встановлення порядку між кластерами [4] довільно обирається один об'єкт з першого кластеру та другий з іншого. Порядок між кластерами встановлюється такий, який прийнято між обраними об'єктами в будь якій з ранжирувань які розглядаються. Раціональність такого підходу доведена в роботі [5]. Можливо виникнення ситуації, коли два об'єкта знаходяться у кластері в усіх наданих ранжируваннях. Внаслідок цього підсумкове ранжирування також описується цим кластером. В результаті фактично одержуємо загальний тренд виконаних робіт або динаміку використання окремої марки рухомого складу. Вона приймає вигляд ранжирування з кластерами, тобто $RS^{h_i} = [t_0 < t_1 = t_j < \dots]$ Формування загальної динаміки призначено для зменшення

впливу шумів або випадкових викидів значень за окремими місяцями на підсумковий результат.

Відсоткова динаміка є змінною величиною. Причиною цьому є сезонні коливання та наявні договори на перевезення. В результаті розрахунки відсоткової динаміки виконуються через визначення середніх величин збільшення (зменшення) за послідовностями які співпадають. Відсотки змін знаходяться як суми між середніми величинами за схожими послідовностями з моменту повторення динаміки. Для пошуку моменту початку однотипної динаміки використовується зворотній перерахунок елементів RS^{h_j} . Здійснюються порівняння передостаннього знаку домінування. В випадку, коли він співпадає – порівнюються попередні знаки. Пошук точки початку завершується, коли знаки не співпадають. Наступним кроком будуються бінарні відношення домінування для всіх місяців

$$BON_{jm}^{h_i} = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & j \end{matrix} \\ \begin{matrix} k \\ k+1 \\ \dots \\ j \end{matrix} & \begin{bmatrix} 1 & bo_{jm}^{h_i}(t_k, t_2) & \dots & bo_{jm}^{h_i}(t_k, t_j) \\ bo_{jm}^{h_i}(t_{k+1}, t_1) & 1 & \dots & bo_{jm}^{h_i}(t_{k+1}, t_j) \\ \dots & \dots & \dots & \dots \\ bo_{jm}^{h_i}(t_j, t_1) & bo_{jm}^{h_i}(t_j, t_2) & \dots & 1 \end{bmatrix} \end{matrix} \quad (6)$$

В подальших розрахунках приймають участь тільки місяці, для яких помноження всіх елементів не дає нуля в обох напрямках. Середня динаміка визначається з рівняння

$$DIN^{h_i} = \sum_{i=kp}^{12} \frac{\sum_{n=tr}^j \frac{(tp^{h_i}(t_n, i) - tp^{h_i}(t_{n+1}, i))}{tp^{h_i}(t_n, i)}}{(tr-1)}, \quad (7)$$

де tr – кількість років після точки початку динаміки, од;

kp – місяці, динаміка яких співпадає з загальним трендом, од.

Одержана динаміка приросту дозволяє визначити обсяг робіт на наступний період по місяцях

$$pt^{h_i}(t_{j+1}, i) = tp^{h_i}(t_j, i) + tp^{h_i}(t_j, i) \cdot DIN^{h_i} \quad i = 1, 2, \dots, 12. \quad (8)$$

Наступним кроком виконується визначення сезонності попиту на окрему марку. Пошук сезонності призначено для аналізу доцільності здачі окремої марки рухомого складу в оренду на відповідний термін. Для цього за даними прогнозу будується бінарне відношення домінування

$$pbo^{h_i}(k, m) = \begin{cases} 1, & \text{якщо } pt^{h_i}(k) \leq pt^{h_i}(m) \\ 0, & \text{якщо } pt^{h_i}(k) > pt^{h_i}(m) \end{cases} \quad k \neq m, \quad k = 1, 2, \dots, 12 \quad m = 1, 2, \dots, 12. \quad (9)$$

У результаті формується матриця, яка описує сезонність перевезень

$$PBO^{h_i} = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & 12 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \dots \\ 12 \end{matrix} & \begin{bmatrix} 1 & pbo^{h_i}(1, 2) & \dots & pbo^{h_i}(1, 12) \\ pbo^{h_i}(2, 1) & 1 & \dots & pbo^{h_i}(2, 12) \\ \dots & \dots & \dots & \dots \\ pbo^{h_i}(12, 1) & pbo^{h_i}(12, 2) & \dots & 1 \end{bmatrix} \end{matrix} \quad (10)$$

Пошук реальної динаміки здійснюється за допомогою протиставлення прогнозним даним результатів попереднього року. За останній рік також будується бінарне відношення домінування (11).

$$sbo^{h_i}(k, m) = \begin{cases} 1, & \text{якщо } tp^{h_i}(t_j, k) \leq tp^{h_i}(t_j, m) \\ 0, & \text{якщо } tp^{h_i}(t_j, k) > tp^{h_i}(t_j, m) \end{cases} \quad k \neq m, \quad k = 1, 2, \dots, 12 \quad m = 1, 2, \dots, 12, \quad (11)$$

$$SBO^{h_i} = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & 12 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \dots \\ 12 \end{matrix} & \begin{bmatrix} 1 & sbo^{h_i}(1, 2) & \dots & sbo^{h_i}(1, 12) \\ sbo^{h_i}(2, 1) & 1 & \dots & sbo^{h_i}(2, 12) \\ \dots & \dots & \dots & \dots \\ sbo^{h_i}(12, 1) & sbo^{h_i}(12, 2) & \dots & 1 \end{bmatrix} \end{matrix}. \quad (12)$$

Аналіз двох відношень (PBO^{h_i} та SBO^{h_i}) дає можливість виділити ядро протиріч. Після його згладжування одержуємо більш точну сезонну динаміку роботи окремої марку рухомого складу тобто сезонну послідовність на наступний рік. Вона приймає вигляд ранжирування з кластерами (однаковий обсяг робіт) від більшого до меншого за обсягами робіт місяця.

Наприклад $PSO^{h_i} = [5 > 6 > \{7, 8\} > 1 > 3 > 4 > 9 > 11 > 10 > 12 > 2]$. Найбільш напруженим можна вважати місяці з 5 по 8 включно. Простими та менш трудомісткими є місяці з 9 по 12 включно. Аналіз обсягів перевезення за ними дозволить прийняти рішення про надання оренди рухомого складу третім особам.

Висновки

В роботі надано новий підхід до прогнозування, який цілком відрізняється від класичних. В ньому прийнято припущення про наявність основного тренду та сезонної складової. Розроблена методика дозволяє визначати точку початку однотипної події. Це дає можливість нівелювати статистичні викиди та похибки. Окремо визначається помісячна динаміка та сезонна складова перевезень. Лише всі результати прогнозування дозволять сформулювати ґрунтовне рішення щодо зміни структури парку рухомого складу у часі.

Доцільним є проведення перевірки запропонованої методики на реальному полігоні транспортного обслуговування. Разом з порівнянням з існуючими моделями це дозволить виділити її слабкі сторони та надати рекомендацій щодо подальшого практичного використання.

Список літератури. 1. Орлов А.И. Прикладная статистика. Учебник // М.: Из-во “Экзамен”, 2004. – 656 с. 2. Вентцель Е.С. Теория вероятностей // М.: Гос изд-во физико-математической литературы, 1962. – 564 с. 3. Демидович Б.П., Марон И.А., Шувалова Э.В. Численные методы анализа // М.: Физматгиз, 1963. – 400 с. 4. Орлов А.И. Экспертные оценки: учебное пособие // М.: Наука, 2002. – 35 с. 5. Горский В.Г., Орлов А.И., Гриценко А.А. Метод согласования кластерных ранжировок // М.: Атоматика и телемеханика №3, 2000. – С.159- 167