

[DOI 10.28925/2663-4023.2025.31.1002](https://doi.org/10.28925/2663-4023.2025.31.1002)

УДК 658.787:519.876.5:004.851

Балвак Андрій Анатолійович

аспірант кафедри комп'ютерної інженерії

Державний університет інформаційно-комунікаційних технологій, Київ, Україна

ORCID: 0000-0002-6441-8225

a.balvak@stud.duikt.edu.ua**Лашевська Наталія Олександрівна**

завідувач кафедри комп'ютерної інженерії

Державний університет інформаційно-комунікаційних технологій, Київ, Україна

ORCID: 0000-0003-2148-115X

n.lashchevska@duikt.edu.ua

МЕТОДИКА ЗБАЛАНСОВАНОГО ПРИЗНАЧЕННЯ МІСЦЬ ЗБЕРІГАННЯ ТОВАРІВ НА ОСНОВІ ЧАСОВИХ ХАРАКТЕРИСТИК ПОПИТУ

Анотація. У статті розглянуто актуальну науково-практичну задачу призначення місць зберігання (SLAP) на складах електронної комерції. Хоча застосування класичних підходів, наприклад, ABC-аналізу дозволяє оптимізувати маршрути комплектування замовлень, але ігноруються операційні витрати на розміщення товарів. Це призводить до дисбалансу, внаслідок чого скорочення відстаней відбору досягається за рахунок збільшення трудомісткості процесу розміщення товарних позицій (SKU). Проведений аналіз останніх досліджень показав, що попри доведену користь окремих методів прогнозування часових рядів, залишається актуальною потреба у комплексних методиках. Зокрема, бракує підходів, які б інтегрували аналіз динаміки попиту безпосередньо у вирішення задачі SLAP та збалансовували пов'язані витрати між процесами розміщення і комплектування замовлень. Запропоновано методику Часово-орієнтованого Призначення Місць Зберігання (TOASL) фундаментальним принципом якої є об'єднання процесів збору замовлень та розміщення товарів у рамках єдиної оптимізаційної задачі. Її архітектура складається з чотирьох ключових фаз. На першому етапі виконується обробка даних, яка включає агрегацію, фільтрацію та нормалізацію часових рядів попиту. На етапі моделювання відбувається кластеризація часових рядів для виявлення груп SKU зі схожою динамікою. Завершують цикл фази симуляції (створення віртуальної моделі складу та тестування політик) та оцінювання (розрахунок ключових показників ефективності й статистична верифікація). В дослідженні, яке охоплює перші два етапи методики, було використано загальнодоступний набір даних з платформи Kaggle про транзакції інтернет-магазину електроніки. Після фільтрації даних, під час якої відібрано події покупок лише для цільових категорій товарів, та агрегації за сесіями, сформовано 471 унікальний часовий ряд попиту SKU. Ці ряди були нормалізовані методом Min-Max для фокусування на формі та динаміці попиту, а не на абсолютних обсягах. Проведено порівняльний аналіз трьох алгоритмів кластеризації: агломеративної ієрархічної кластеризації (АНС) за методом Уорда (Ward), k-середніх (k-means) та самоорганізаційних карт (SOM). Оптимальну кількість кластерів, яка дорівнює 10, визначено методом ліктя. Найкращу якість розбиття продемонстрував метод АНС-Ward, досягнувши найвищого коефіцієнта силуету зі значенням 0,39. Виконане групування дозволило виявити 10 наборів позицій зі схожою динамікою. На основі аналізу 5 найбільших кластерів запропоновано логічно обґрунтовані принципи зонування складу, що передбачають компактне розміщення товарів зі схожим попитом у суміжних проходах. Доведено, що кластеризація часових рядів попиту є ефективною основою для подальшого адаптивного зонування складу.

Ключові слова: призначення місць зберігання; кластеризація часових рядів; динаміка попиту; складська логістика; комплектування замовлень; розміщення товарів; електронна комерція; симуляція; обробка даних; алгоритм; прогнозування.



ВСТУП

Постановка проблеми. Задача призначення місць зберігання полягає у такому розміщенні товарних позицій (Stock Keeping Unit, SKU), щоб мінімізувати сумарну довжину та час маршрутів під час комплектування замовлень з урахуванням фактичного попиту. Водночас слід дотримуватися операційних обмежень складу, зокрема вимог до планування і геометрії, розташування проходів і точок приймання та відвантаження, місткості зон, а також базових правил безпеки й сумісності товарів. У загальному випадку це NP-складна комбінаторна оптимізація, що потребує практично придатних евристичних та аналітичних підходів для великих асортиментів товарів і динамічних умов електронної комерції. Класичні політики, зокрема ABC-аналіз, зосереджуються переважно на ефективності збирання товарів. У межах ABC номенклатуру поділяють на класи А, В, С за внеском у попит чи оборот (за правилом Парето) і для товарів класу А резервують пріоритетні високодоступні зони зберігання. Однак така схема зазвичай не враховує витрати й навантаження на розміщення та поповнення, зокрема збільшення довжини маршрутів від місць приймання до місць зберігання. У підсумку виникає суперечність критеріїв, коли схема оптимальна за мінімізацією відстані під час комплектування, супроводжується істотним зростанням витрат і часу операцій при розміщенні товарів. Отже, науково-практичне завдання дослідження – обґрунтувати та перевірити підхід до часово орієнтованого призначення місць зберігання. Він має спиратися на кластеризацію часових рядів попиту, оптимізувати маршрути комплектування та збалансувати витрати між процесами розміщення і відбору.

Аналіз останніх досліджень і публікацій. Наукові праці з призначення місць зберігання традиційно зосереджуються на правилах класового розміщення, аналізі кошика покупок (асоціативні правила) та на евристичних для організації складу й маршрутів комплектувальника. Зазвичай ідеться про політики розміщення, такі як «найближче вільне місце» та «виділене розміщення», а також про правила маршрутизації, застосування яких дає змогу зменшити довжину маршруту під час комплектування товарів. Автори [1] представляють алгоритм на основі інтелектуального аналізу даних з використанням асоціативних правил для виявлення взаємопов'язаних товарів та їх спільного розміщення на близьких місцях зберігання. Вони безпосередньо мінімізують сумарну відстань двох процесів складування та відбору замовлень і демонструють переваги над базовими політиками адресації місць зберігання. Водночас у роботі не моделюється часова мінливість попиту та не розглядаються часові ряди як самостійний об'єкт аналізу, що обмежує застосовність підходу в умовах сезонності та трендів.

У дослідженні [2] описано гібридну стратегію призначення місць зберігання на основі кластеризації часових рядів попиту та відібрано найкращу модель за коефіцієнтом силуету. На наборі даних історії замовлень інтернет-магазину з платформи Kaggle автори продемонстрували помітні вигоди за довжиною маршруту комплектування для політик S-подібного маршруту та маршруту з серединою проходу, порівняно з випадковим розміщенням і ABC. Однак низка припущень робить формалізацію задачі спрощеною, адже на складі відсутні обмеження місткості та сумісності, передбачено одну точку входу-виходу і застосовано моноблокове планування без поперечних проходів. Розміщення вважається незмінним у проміжках між періодичними оновленнями плану зонування й адресації місць зберігання. Метрики якості зосереджені переважно на маршрутах комплектування, тоді як трудомісткість і протяжність переміщень для розміщення товарів на зберігання окремо не вимірюються й не аналізуються.



Крім того, у роботі [3] запропоновано прогнозування на найближчий тиждень (шість робочих днів) щоденного навантаження на складські ділянки приймання, зберігання, комплектування та пакування. Автори застосовують моделі часових рядів і XGBoost, попередньо перевіряючи стаціонарність даних тестом Kwiatkowski–Phillips–Schmidt–Shin, який виявляє наявність тренду та нестабільності. Автори порівнюють ARMA/ARIMA/ARIMAX (класичні лінійні моделі з можливими зовнішніми змінними), а також експоненційне згладжування та бустинг. Точність оцінюється за RMSE, MAE, MAPE (середньоквадратична, середня абсолютна та відносна відсоткова помилка). Показано, що результати істотно різняться між ділянками, причому пакування прогнозується найкраще, тоді як зберігання – найскладніше. Важливими пояснювальними ознаками виявилися значення показників, зміщені в часі (значення попередніх днів), а також індикатори аномальних днів. Але автори не моделюють сезонні компоненти, а використання SARIMA чи SARIMAX лише позначено як напрям подальших досліджень. Фокус зроблено на прогнозуванні обсягу операцій на складських ділянках, а не на інтеграції цих прогнозів у призначення місць зберігання чи оптимізацію маршрутів комплектування. Витрати та відстані при розміщенні товарів на полицях не оцінюються. Отже, робота важлива як доказ користі прогнозів часових рядів для планування персоналу та вирівнювання навантаження. Однак залишає відкритим питання використання цих прогнозів для часово чутливого SLAP і збалансування витрат між процесами зберігання та відбору.

Диференційований підхід до прогнозування в e-commerce представлено у статті [4], де ARIMA використано для помісячних рівнів запасів, а LSTM – для щоденних продажів по 350 категоріях за три місяці. Автори показують, що така «подвійна» схема підвищує точність і може слугувати основою для зонування складу та планування ресурсів. При цьому модель є одномірною (уніваріантною) і не інтегрує прогнози безпосередньо в призначення місць зберігання чи оцінювання відстаней і трудомісткості розміщення та відбору. Як подальший крок автори окреслюють врахування багатофакторних взаємодій. Отже, робота підсилює аргументацію на користь часових рядів у складській логістиці, але залишає відкритим питання їх поєднання зі SLAP і збалансування витрат між процесами, що й становить фокус даного дослідження.

У дослідженні [5] описано прогностичну модель, орієнтовану на дані для прогнозування кількості позицій до відвантаження на п'ять робочих днів наперед з метою оперативного планування чисельності й графіків змін персоналу. Модель, апробована на кейсі однієї компанії, перевершує суб'єктивні ручні прогнози фахівців та підходить на основі автоматизованого машинного навчання за точністю короткострокових прогнозів і дає змогу зменшити ризики недостатнього та надмірного укомплектування змін. Проте в роботі не інтегровано часові прогнози у призначення місць зберігання та не оцінено відстані і витрати процесів розміщення й відбору. Автори окремо зазначають обмеження узагальнюваності результатів, адже дослідження виконано на даних одного підприємства. Також наголошено на потребі перейти від прогнозу кількості позицій до відвантаження до прогнозу кількості операцій відбору, що дозволить точніше нормувати працю. Отже, стаття підкріплює користь часових рядів для управління складським навантаженням, але залишає відкритим зв'язок із SLAP і балансуванням витрат між процесами.

В роботі [6] проаналізовано інтелектуальні підходи до вирішення задач призначення місць зберігання, комплектування замовлень та впровадження роботизованих мобільних систем. Автори виокремили як одні з найбільш перспективних напрямів алгоритми, що поєднують аналіз історичних даних, кореляційні зв'язки між



товарами та, зокрема, кластеризацію часових рядів. Водночас підкреслено, що залишаються недостатньо дослідженими питання адаптивної кластеризації з урахуванням сезонності та інтеграції прогнозних моделей, наприклад, на основі довгої короткочасної пам'яті у динамічне управління розміщенням.

Отже, проведений аналіз досліджень і публікацій виявляє суттєву наукову прогалину. Попри доведену користь окремих методів прогнозування для управління персоналом чи рівнями запасів, наявні дослідження демонструють брак комплексних підходів. Зокрема, бракує методик, які б інтегрували аналіз часових рядів попиту безпосередньо у вирішення задачі призначення місць зберігання та збалансування пов'язаних із цим операційних витрат на процеси розміщення та відбору.

Мета статті. Метою даного дослідження є розробка та часткова апробація комплексної методики часово-орієнтованого призначення місць зберігання, спрямованої на вирішення науково-практичної задачі збалансування операційних витрат між процесами розміщення товарів та комплектування замовлень. На відміну від статичних підходів, у роботі увагу зосереджено на глибокому аналізі динаміки попиту. Тому основними завданнями дослідження є обґрунтування архітектури методики, що базується на кластеризації часових рядів попиту SKU, а також проведення порівняльного аналізу ефективності алгоритмів кластеризації на реальних даних інтернет-магазину для виявлення груп товарів зі схожою динамікою. Кінцевим завданням є доведення, що таке групування слугує ефективною та логічно обґрунтованою основою для подальшого адаптивного зонування складу.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Для досягнення поставлених у дослідженні цілей пропонується авторська методика, названа Часово-орієнтованим Призначенням Місць Зберігання (Time-Oriented Assignment of Storage Locations, TOASL). На відміну від поширених статичних стратегій, таких як ABC-аналіз, що оперує усередненими показниками чи частотою звернень, її особливість полягає у глибокому аналізі саме динаміки попиту для кожної товарної позиції. TOASL дозволяє враховувати специфічні для номенклатури часові закономірності, такі як періодичні цикли, сезонні зміни та довгострокові тренди [7]. Ключовою метою методики є знаходження оптимального розташування товарів для мінімізації сукупних операційних витрат, збалансовуючи ефективність як процесу збору замовлень (picking), так і процесу розміщення товарів на зберігання (put-away).

В основі методики TOASL лежить об'єднання процесів збору замовлень та розміщення товарів у рамках єдиної оптимізаційної задачі. Як свідчить аналіз існуючих практик управління складом, зосередження зусиль лише на скороченні шляху під час комплектування часто не враховує пов'язані витрати на розміщення новоприбулих товарів. Запропонований підхід запобігає виникненню дисбалансу, коли оптимізація процесу відбору досягається за рахунок суттєвого збільшення трудомісткості розміщення.

Щоб досягти загальної операційної ефективності, TOASL використовує зважену цільову функцію Z , яка спрямована на мінімізацію сукупної пройденої відстані для обох ключових процесів:

$$Z = \alpha \cdot TotalPickingDistance + \beta \cdot TotalPutawayDistance \quad (1)$$



У цій формулі TotalPickingDistance представляє загальну відстань, пройдену під час комплектування всіх замовлень за певний часовий проміжок, TotalPutawayDistance – загальну відстань, пов'язану з розміщенням товарів на зберігання. Коефіцієнти α і β слугують вагами, які визначають відносну важливість кожного з цих процесів для конкретного підприємства. Саме вибір значень α та β встановлює стратегічний баланс між швидкістю виконання замовлень та ефективністю розміщення. Наприклад, якщо пріоритетом є максимальна швидкість комплектування, доцільно обрати $\alpha > \beta$. Визначення конкретних значень цих вагових коефіцієнтів може базуватися на експертній оцінці керівництва складу або ж на основі детального аналізу історичних даних щодо вартості та тривалості відповідних операцій.

Для детальнішого представлення, розглянемо цю комбіновану цільову функцію в контексті конкретного набору замовлень $O = \{O_1, \dots, O_Z\}$ та асортименту товарів $P = \{p_1, \dots, p_N\}$. Нехай план розміщення χ визначає місце зберігання для кожного товару p_i . Тоді функція Z набуває вигляду:

$$Z = \alpha \cdot \sum_{z=1}^Z D_{pick}(O_z, \chi) + \beta \cdot \sum_{i=1}^N D_{put}(p_i, \chi), \quad (2)$$

де $D_{pick}(O_z, \chi)$ – відстань (або час) для виконання (комплектуювання) одного замовлення O_z при плані розміщення χ . Відповідно, $D_{put}(p_i, \chi)$ представляє відстань, яку потрібно подолати для розміщення товару p_i на його призначене місце зберігання згідно з планом χ . Використання такого критерію гарантує, що оптимізація одного процесу, наприклад, збору, не призведе до небажаного збільшення витрат при розміщенні, забезпечуючи збалансовану ефективність складських операцій [8].

Архітектура методики TOASL

Архітектура методики TOASL передбачає послідовну реалізацію чотирьох ключових фаз, що забезпечують повний цикл аналізу від вхідних даних до оцінки ефективності розміщення.

Перший етап підготовки даних включає збір та очищення транзакцій, що передбачає усунення некоректних записів, таких як аномалії чи пропуски. Далі відбувається відбір пріоритетних товарних позицій, для чого може застосовуватися критерій Парето, та формування нормованих часових рядів попиту, приведених до єдиного масштабу від 0 до 1, для подальшого аналізу.

На другому етапі моделювання застосовують алгоритми кластеризації часових рядів для виявлення груп товарів зі схожою динамікою попиту [9]. Серед них, як правило, розглядають агломеративну ієрархічну кластеризацію (АНС), алгоритм кластеризації методом k-середніх (k-means) та алгоритм самоорганізаційних карт (SOM). Оптимальна кількість кластерів визначається за допомогою пошуку характерної точки зламу на діаграмі або аналізом дендрограми, а якість розбиття оцінюється за коефіцієнтом силуєту. Результатом є призначення кожного кластера SKU певній зоні складу та, за потреби, внутрішньозонава оптимізація розміщення за інтенсивністю попиту.

Третій етап, симуляція, передбачає створення віртуальної моделі складу, що враховує його фізичну конфігурацію таку як кількість проходів та розташування точки входу/виходу. У цьому середовищі тестуються різні стратегії комплектування та політики маршрутизації, такі як S-подібний, маршрут з серединою проходу, маршрут найбільшого розриву та інші. Важливою складовою є розрахунок відстаней як для

процесів збору замовлень, так і розміщення товарів, що дозволяє оцінювати збалансовані плани зберігання [10].

Завершальний етап, оцінювання, охоплює розрахунок ключових показників ефективності, зокрема середньої та сумарної довжини маршрутів для різних сценаріїв. Для перевірки статистичної значущості відмінностей між політиками розміщення застосовується однофакторний дисперсійний аналіз (ANOVA) та post-hoc тести, зокрема критерій Тьюкі HSD, для попарного порівняння.

Загальний порядок проведення експериментів відповідно до запропонованої архітектури наведено на рис. 1.

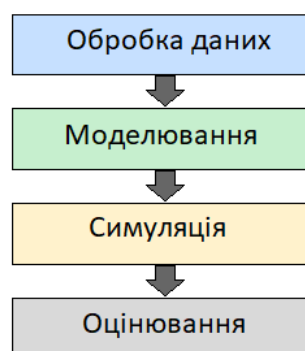


Рис. 1. Етапи експериментів згідно з методикою TOASL

Така структурована послідовність кроків забезпечує системний підхід до розробки та валідації часово-орієнтованих стратегій розміщення товарів на складі.

Опис та підготовка даних

Для експериментальної перевірки запропонованої методики TOASL було використано загальнодоступний набір даних eCommerce events history in electronics store, розміщений на платформі Kaggle [11]. Цей дата сет містить дані про 885129 подій, зафіксованих у великому інтернет-магазині електроніки за період з вересня 2020 р. по лютий 2021 р.

Кожен запис у наборі даних фіксує одну дію користувача та включає дев'ять характеристик:

- event_time – це часова мітка події;
- event_type – тип події, що може приймати значення view (перегляд), cart (додавання в кошик) або purchase (покупка);
- product_id – ідентифікатор продукту;
- category_id – ідентифікатор категорії продукту;
- category_code – текстовий код категорії (наприклад, 'electronics.telephone');
- brand – назва бренду;
- price – ціна продукту;
- user_id – ідентифікатор користувача;
- user_session – тимчасовий ідентифікатор сесії користувача, який змінюється щоразу, коли клієнт повертається після тривалої паузи.

У даному дослідженні, на відміну від інших робіт в яких проаналізовано весь набір подій, було сформовано вузьку цільову вибірку. Для побудови часових рядів фактичного попиту враховувалися лише події типу event_type = purchase. Крім того, для забезпечення однорідності асортименту та моделювання розміщення фізично сумісних товарів,



вибірку було обмежено лише двома основними категоріями, для яких значення в полі `category_code` починаються зі слів `computers.` та `electronics.` Ці категорії охоплюють відносно негабаритні товари, зокрема смартфони, ноутбуки, аксесуари й комп'ютерні компоненти, що робить їх зручними для спільного розміщення у зонах стележного зберігання.

event_time	event_type	product_id	category_id	category_code	brand	price	user_id	user_session
2020-09-24 12:04:10 UTC	purchase	1507291	2,14442E+18	computers.components.power_supply	supermicro	217.57	1,51592E+18	xn6SHCnZtk
2020-09-24 12:15:06 UTC	purchase	822426	2,14442E+18	computers.peripherals.camera	logitech	123.35	1,51592E+18	2gngxS29Ts
2020-09-24 12:19:01 UTC	purchase	4060928	2,14442E+18	electronics.video.tv		89.22	1,51592E+18	3yFCkx2KKW
2020-09-24 12:25:18 UTC	purchase	4060928	2,14442E+18	electronics.video.tv		89.22	1,51592E+18	3yFCkx2KKW
2020-09-24 12:29:49 UTC	purchase	1283197	2,14442E+18	computers.peripherals.nas	zyxel	123.79	1,51592E+18	3jFpdbozOd
2020-09-24 12:29:49 UTC	purchase	888783	2,14442E+18	computers.components.hdd		83.02	1,51592E+18	3jFpdbozOd
2020-09-24 12:39:40 UTC	purchase	1674454	2,14442E+18	electronics.telephone		11.86	1,51592E+18	v1o5QvC6hH

Рис. 2. Фрагмент даних, відібраних для аналізу попиту

Перш ніж перейти до детального опису застосованої процедури, варто зазначити, що набір даних `eCommerce events history in electronics store` вже використовувався у наукових дослідженнях, однак для вирішення інших завдань. Наприклад, у роботі [12] ці дані були застосовані для виявлення аномалій у поведінці користувачів. Дослідження зосереджено на застосуванні методів машинного навчання, зокрема методу опорних векторів, для ідентифікації нетипових дій, які можуть свідчити про шахрайство або збої в системі. Автори публікації підійшли до завдання шляхом перетворення його на бінарну класифікацію. Аномальними вважалися ті події, цінний показник яких значно відрізнявся від нормального розподілу. А саме: порогове значення було встановлено на рівні 95-го перцентилі й будь-яка подія, пов'язана з продуктом, ціна якого була вищою за цей поріг, маркувалася як аномалія, а всі інші – як нормальні.

В іншому дослідженні [13] набір даних `eCommerce events history in electronics store` слугував основою для прогнозування факту покупки. Автори використовували методи послідовного аналізу шаблонів (`Sequential Pattern Mining, SPM`) для дослідження поведінки користувачів. З метою виділення ознак (найбільш частих послідовностей) порівняно три різні алгоритми `SPM`: `PrefixSpan`, що є методом, який базується на рості префікс-проекцій; `ClaSP`, орієнтований на пошук замкнених послідовних шаблонів; `VMSP`, що реалізує метод вертикального видобутку максимальних шаблонів. Отримані таким чином ознаки подавалися на вхід двох моделей машинного навчання для класифікації, а саме: `Extreme Gradient Boosting`, що є ансамблевим методом на основі дерев рішень та `k-nearest neighbor`, який є метричним класифікатором, що базується на відстані до найближчих сусідів.

Мета даного дослідження принципово відрізняється й полягає не у виявленні аномалій чи прогнозуванні факту покупки, а в аналізі підтверджених даних про транзакції з обраних категорій для кластеризації часових рядів попиту на товари з метою їх подальшої складської оптимізації за методикою `TOASL`.

Моделювання та підготовка даних виконувалися у хмарному середовищі `Google Colab`. Аналіз реалізовано мовою програмування `Python` із залученням спеціалізованих бібліотек. Для ефективної агрегації транзакцій, групування сесій та формування структур даних використовувалася бібліотека `pandas`, тоді як `scikit-learn` була застосована на етапі попередньої обробки, зокрема для нормалізації часових рядів методом `Min-Max`.

Процес обробки відфільтрованих даних, що охоплюють лише події `purchase` для категорій `computers.*` та `electronics.*`, включав наступні кроки. По-перше, ці транзакції



було згруповано за ідентифікатором сесії користувача (`user_session`) для виявлення покупок, що мали декілька різних товарів. Для кожної сесії було визначено час її завершення, який відповідав максимальній часовій мітці події, та підраховано кількість замовлених унікальних товарів. На рис. 3 наведено приклад таких даних, де відфільтровано сесії, що включають два або більше артикулів.

user_session	product_id	session_end_time	purchased_quantity	event_type	category_code	brand	price
3JFpdbozOd	888783	2020-09-24 12:29:49	1	purchase	computers.components.hdd		83,02
3JFpdbozOd	1283197	2020-09-24 12:29:49	1	purchase	computers.peripherals.nas	zyxel	123,79
UL6gfSiAth	3648406	2020-09-24 19:18:00	1	purchase	computers.components.videocards	asus	1450,13
UL6gfSiAth	3758770	2020-09-24 19:18:00	1	purchase	computers.components.motherboard	asus	146,4
UL6gfSiAth	4013167	2020-09-24 19:18:00	1	purchase	electronics.telephone	samsung	1216,24
UL6gfSiAth	4154233	2020-09-24 19:18:00	1	purchase	computers.components.cpu	intel	533,56
TCogRQubOO	1578301	2020-09-25 13:24:11	1	purchase	computers.components.motherboard	gigabyte	77,06
TCogRQubOO	3791403	2020-09-25 13:24:11	1	purchase	computers.components.cpu	amd	236,97

Рис. 3. Фрагмент даних, що ілюструє агрегацію багатопозиційних сесій

На наступному етапі на основі часу завершення сесій було побудовано матрицю добового попиту для кожного унікального `product_id`, в результаті чого сформовано фінальну вибірку з 471 унікальної товарної позиції. Для підготовки цих даних до класифікації, часові ряди попиту для кожного товару було нормалізовано у діапазон $[0,1]$ за методом Min-Max. Ця процедура лінійного масштабування приводить найнижчий попит до нуля, а максимальний пік попиту для цього товару – до 1. Такий підхід усуває вплив абсолютних обсягів продажів, дозволяючи алгоритмам кластеризації фокусуватися виключно на формі та динаміці попиту.

Кластеризація часових рядів попиту SKU

Різноманітність динаміки у 471 нормалізованому часовому ряді попиту для SKU створює передумови для їх кластеризації. Метою цього етапу є виявлення груп товарів зі схожими часовими характеристиками, таких як товари зі стабільним попитом, сезонними піками або нерегулярними продажами, для їх подальшого спільного розміщення на складі.

Для вирішення цього завдання було проведено порівняльний аналіз трьох поширених алгоритмів кластеризації, кожен з яких має різну методологію групування.

Агломеративна ієрархічна кластеризація. Це об'єднувальний метод, що працює за висхідним принципом, де на початковому етапі кожен об'єкт, у даному випадку – часовий ряд SKU, розглядається як окремий кластер. На кожному наступному кроці поєднуються дві найближчі групи, доки всі об'єкти не утворять єдину спільноту з чіткою ієрархією. Для визначення «близькості» було обрано метод Уорда (Ward's method). Ця стратегія ґрунтується на мінімізації приросту загальної внутрішньокластерної дисперсії (WCSS) при кожному об'єднанні. Такий підхід спрямований на формування компактних, сферичних кластерів, оскільки алгоритм надає перевагу об'єднанням, що найменше збільшують загальну дисперсію, та уникає злиття віддалених груп.

к-середніх. Це ітераційний алгоритм кластеризації, що розподіляє дані на заздалегідь визначену кількість кластерів k . Його робота полягає у двох основних кроках, що ітеративно повторюються. Перший крок – це етап призначення, на якому кожна точка даних SKU відноситься до кластера з найближчим центроїдом. Ця близькість зазвичай визначається за евклідовою відстанню. Другий крок – це етап оновлення, на якому центроїди кожного кластера перераховуються як середнє арифметичне всіх точок, що

увійшли до нього. Ітеративне виконання цих кроків продовжується до збіжності, аби мінімізувати WCSS (суму квадратів відстаней від точок до центроїдів їхніх кластерів).

Самоорганізаційні карти. Карти Кохонена є методом некерованого навчання, що базується на архітектурі нейронної мережі. Алгоритм SOM здійснює проектування багатовимірних вхідних даних (часових рядів) на низькорозмірну, зазвичай двовимірну, сітку (карту) нейронів. У процесі навчання для кожного вхідного вектора, шляхом конкурентного процесу, алгоритм визначає нейрон, ваговий вектор якого є найближчим до цього вхідного вектора. Визначений найкращий вузол та його сусіди на карті оновлюють свої вагові вектори, наближуючи їх до вхідного вектора. Таким чином, SOM не просто кластеризує дані, а й зберігає їх топологічні зв'язки, оскільки схожі часові ряди активують нейрони, що розташовані близько один до одного на карті.

Ключовим параметром для алгоритмів k-means та SOM є попереднє визначення оптимальної кількості кластерів k . Її було встановлено за допомогою методу ліктя, який ґрунтується на аналізі залежності WCSS від k . Графік на рис. 4 демонструє виразний злам («лікоть») при $k=10$, що вказує на цю кількість як на найбільш збалансований вибір, який забезпечує достатню деталізацію без надмірної фрагментації.

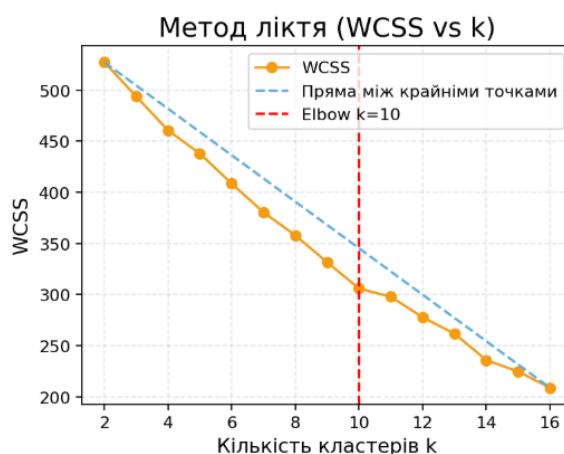


Рис. 4. Визначення оптимальної кількості кластерів методом ліктя

Для підсумкового порівняння якості розбиття за $k=10$ використано коефіцієнт силуету. Це безрозмірна міра в інтервалі від -1 до 1 , яка показує наскільки об'єкти ближчі до «свого» кластера і віддалені від сусідніх; що вище значення, то краща якість групування. Найкращий результат показала агломеративна ієрархічна кластеризація з критерієм Уорда і значенням $0,39$ (табл. 1). Отже, АНС-Ward приймається як базовий підхід для етапу кластеризації, що передуює призначенню місць зберігання за методикою TOASL.

Таблиця 1

Порівняння алгоритмів кластеризації часових рядів ($k=10$)

Алгоритм	Міра/Критерій зв'язування	Коефіцієнт силуету
АНС	Метод Уорда	0,39
SOM	Евклідова відстань	0,12
k-means	Евклідова відстань	0,03

Отримане значення $k=10$ було верифіковано за допомогою ієрархічної дендрограми, побудованої за методом АНС-Ward (рис. 5). Встановлення горизонтального порогу, позначеного пунктирною лінією, на рівні, який відповідає поділу на 10 кластерів, підтверджує чітке та природне розділення груп.

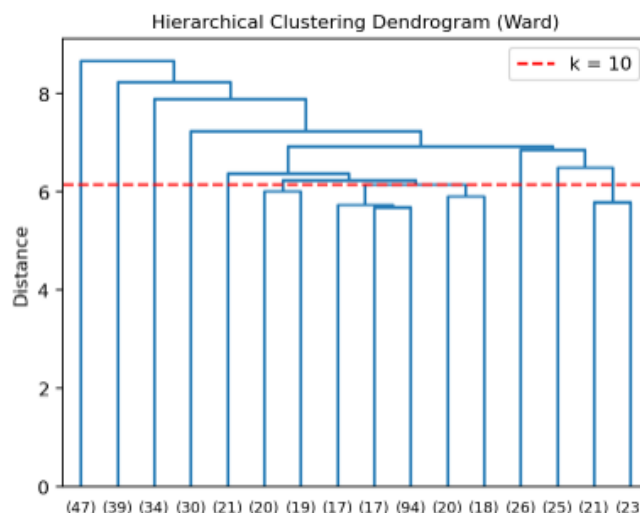


Рис. 5. Ієрархічна дендрограма кластеризації 471 SKU

Загалом було отримано 10 кластерів (з 47, 39, 34, 30, 21, 39, 166, 26, 25 та 44 SKU відповідно). Для формування зонування складу доцільно зосередитись на п'яти найвиразніших групах. У кластері 7 налічується 166 SKU, що формують ядро комп'ютерних комплектуючих. У ньому переважають материнські плати, відеокарти, процесори, блоки живлення, жорсткі диски, а також кулери. Тому групу доцільно розмістити в найближчій до відвантаження зоні А та організувати послідовне розташування товарів – спочатку процесори й кулери, далі блоки живлення та материнські плати, а відеокарти розмістити на суміжних полицях у тому ж проході. Такий порядок скорочує шлях відбору під час збирання замовлень на комплектуючі для персональних комп'ютерів. У першому кластері із 47 SKU переважають відеокарти. Відповідно, такі товари слід розмістити безпосередньо поряд із зоною А на середньому ярусі полиць, щоб скоротити час відбору поширених наборів замовлень, коли спочатку беруть відеокарту, а потім материнську плату та процесор. Кластер 10 (44 SKU) має змішаний склад із блоків живлення, відеокарт, материнських плат та жорстких дисків. Його доцільно розмістити у сусідньому проході поруч із кластером 7 і кластером 1, що скорочує кількість кроків відбору між найчастішими комбінаціями товарів. Кластер 2 із 39 SKU містить значну частку смартфонів разом із принтерами та окремими компонентами. Тому його доцільно розмістити в зоні В ближче до пакування для швидкого відбору смартфонів і аксесуарів, водночас зберігши близькість до основної зони компонентів, щоб забезпечити короткий доступ до блоків живлення та карт пам'яті. Нарешті, кластер 3 (34 SKU) концентрує проектори й принтери, тобто відносно габаритні та крихкі товари. Їх доцільно згрупувати в зоні С з ширшими проходами і нижчими полицями, щоб зручніше розміщувати великі коробки та рідше повертатися до тих самих полиць під час відбору. Таким чином, доцільно компактно розмістити кластери 1, 7 та 10 з комп'ютерними комплектуючими у суміжних проходах однієї зони, з мінімальними



відстанями між полицями. Кластер смартфонів та аксесуарів слід винести ближче до зони пакування, а групу проєкторів і принтерів розмістити у секторі з ширшими проходами та нижчими полицями. Таке планування скорочує довжину маршрутів, зменшує перетини потоків і підвищує швидкість відбору без погіршення заповнюваності стелажів. Інші кластери можуть виконувати роль буферів і прилягати до найближчих за профілем зон [16].

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У результаті проведеного дослідження було обґрунтовано та запропоновано авторську методику Часово-орієнтованого Призначення Місць Зберігання (TOASL). На відміну від класичних підходів, вона спрямована на вирішення операційного дисбалансу між ефективністю процесів збору замовлень та розміщення товарів на зберігання. Для цього запропоновано комбінований критерій оптимізації та архітектуру, що інтегрує аналіз часових рядів попиту [7], [14].

У межах роботи було здійснено підготовку та обробку даних на основі реальних транзакцій інтернет-магазину, в результаті чого сформовано 471 унікальний часовий ряд попиту SKU. Проведено порівняльний аналіз алгоритмів кластеризації, зокрема АНС, k-means та SOM та доведено доцільність застосування агломеративної ієрархічної кластеризації за методом Уорда. Цей метод продемонстрував найкращу якість розбиття за коефіцієнтом силуету та дозволив виявити 10 гомогенних кластерів товарів, таких як проєктори, принтери, смартфони й комп'ютерні компоненти, які мають схожу динаміку попиту. Таким чином, результати дослідження доводять, що кластеризація за часовою динамікою є ефективною та логічно обґрунтованою основою для подальшого зонування складу [15].

В статті охоплено етапи підготовки даних та моделювання в рамках архітектури TOASL. Подальші дослідження мають бути спрямовані на повноцінну реалізацію етапів симуляції та оцінювання. Першочерговим завданням є розробка імітаційної моделі складу, що включатиме формалізацію його геометрії, зокрема конфігурації проходів, розташування точки входу/виходу, та правил розрахунку відстаней. Наступним кроком стане програмна реалізація та експериментальне порівняння різних політик маршрутизації комплектувальника, таких як S-подібна, з серединою проходу та найбільшого розриву. Окремо необхідно дослідити ефективність підходів до відбору товарів, а саме: комплектування одиничних замовлень та партій, кластерне й зональне комплектування при застосуванні часово-орієнтованих та класичних планів розміщення. Фінальна статистична верифікація ефективності методики TOASL потребуватиме застосування дисперсійного аналізу ANOVA та критерію Тьюкі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Pang, K.-W., & Chan, H.-L. (2016). Data mining-based algorithm for storage location assignment in a randomised warehouse. *International Journal of Production Research*, 55(14), 4035–4052. <https://doi.org/10.1080/00207543.2016.1244615>
2. Kalkha, H., et al. (2024). Enhancing warehouse efficiency with time series clustering: A hybrid storage location assignment strategy. *IEEE Access*, 1. <https://doi.org/10.1109/access.2024.3386887>
3. Kalafat, İ., et al. (2021). Workload forecasting of warehouse stations: Comparison between classical time series methods and XGBoost. *Data Science and Applications*, 4(2), 19–24.



4. Wang, C., & Wang, J. (2025). Research on e-commerce inventory sales forecasting model based on ARIMA and LSTM algorithm. *Mathematics*, 13(11), 1838. <https://doi.org/10.3390/math13111838>
5. Eichenseer, P., Hans, L., & Winkler, H. (2025). A data-driven machine learning model for forecasting delivery positions in logistics for workforce planning. *Supply Chain Analytics*, 9, 100099. <https://doi.org/10.1016/j.sca.2024.100099>
6. Balvak, A., & Lashchevska, N. (2025). Intelligent approaches for enhancing warehouse efficiency: Item placement, order picking, and robotic automation. *Cybersecurity: Education, Science, Technique*, 1(29), 161–177. <https://doi.org/10.28925/2663-4023.2025.29.869>
7. Islam, M. S., & Uddin, M. K. (2023). Correlated storage assignment approach in warehouses: A systematic literature review. *Journal of Industrial Engineering and Management*, 16(2), 294. <https://doi.org/10.3926/jiem.4850>
8. Park, J., Park, C., & Hong, S. (2023). Gaussian process-based storage location assignments with risk assessments for progressive zone picking systems. *Computers & Industrial Engineering*, 109700. <https://doi.org/10.1016/j.cie.2023.109700>
9. Keung, K. L., Lee, C. K. M., & Ji, P. (2021). Data-driven order correlation pattern and storage location assignment in robotic mobile fulfillment and process automation system. *Advanced Engineering Informatics*, 50, 101369. <https://doi.org/10.1016/j.aei.2021.101369>
10. Xu, C., Zhao, M., & Li, H. (2024). Data-driven simulation methodology for exploring optimal storage location assignment scheme in warehouses. *Computers & Industrial Engineering*, 110627. <https://doi.org/10.1016/j.cie.2024.110627>
11. Kaggle. (n.d.). *eCommerce events history in electronics store*. <https://www.kaggle.com/datasets/mkechinov/e-commerce-events-history-in-electronics-store>
12. Bollu, S. S. (2024). *Anomaly detection of user behavioral events in e-commerce electronics stores using SVMs* (Bachelor's thesis). Blekinge Institute of Technology.
13. Van Heesch, R. J. A. W. M. (2022). *Predicting a customer's e-commerce purchase: A study that combines sequential pattern mining with machine learning* (Master's thesis). Tilburg University.
14. Zarinchang, A., et al. (2023). Adaptive warehouse storage location assignment with considerations to order-picking efficiency and worker safety. *Journal of Industrial and Production Engineering*, 1–20. <https://doi.org/10.1080/21681015.2023.2263009>
15. Zhang, R.-Q., Wang, M., & Pan, X. (2019). New model of the storage location assignment problem considering demand correlation pattern. *Computers & Industrial Engineering*, 129, 210–219. <https://doi.org/10.1016/j.cie.2019.01.027>
16. Abramov, V., Astafieva, M., Boiko, M., Bodnenko, D., Bushma, A., Vember, V., Hlushak, O., Zhyltsov, O., Ilich, L., Kobets, N., Kovaliuk, T., Kuchakovska, H., Lytvyn, O., Lytvyn, P., Mashkina, I., Morze, N., Nosenko, T., Proshkin, V., Radchenko, S., & Yaskevych, V. (2021). Theoretical and practical aspects of the use of mathematical methods and information technology in education and science. <https://doi.org/10.28925/9720213284km>

**Andrii Balvak**

PhD student at the Department of Computer Engineering
State University of Information and Communication Technologies, Kyiv, Ukraine
ORCID: 0000-0002-6441-8225
a.balvak@stud.duikt.edu.ua

Natalia Lashchevska

Head of the Department of Computer Engineering
State University of Information and Communication Technologies, Kyiv, Ukraine
ORCID: 0000-0003-2148-115X
n.lashchevska@duikt.edu.ua

METHODOLOGY OF BALANCED ASSIGNMENT OF STORAGE PLACES FOR GOODS BASED ON TIME CHARACTERISTICS OF DEMAND

Abstract. The article considers the current scientific and practical problem of storage place assignment (SLAP) in e-commerce warehouses. Classical approaches, such as ABC analysis, optimise order picking routes while ignoring the operational costs of placing goods. This leads to an imbalance, resulting in the reduction of picking distances by increasing the labor intensity of the product put-away process. The analysis of recent studies has shown that despite the proven benefits of individual time series forecasting methods, the need for comprehensive methods remains relevant. In particular, there is a lack of approaches that would integrate the analysis of demand dynamics directly into solving the SLAP problem and balance the associated costs between the processes of placing and picking orders. The Time-Oriented Assignment of Storage Locations (TOASL) methodology is proposed, the fundamental principle of which is to combine the processes of order picking and goods placement within a single optimization problem. Its architecture consists of four key phases. The first stage involves data processing, which includes aggregation, filtering, and normalisation of demand time series. The modelling stage includes clustering of time series to identify groups of SKUs with similar dynamics. The simulation phase (creation of a virtual warehouse model and testing policies) and evaluation (calculation of key performance indicators and statistical verification) complete the cycle. The study, which covers the first two stages of the methodology, used a publicly available dataset from the Kaggle platform on transactions of an online electronics store. After data filtering, which selected only purchase events for target product categories, and aggregation by session, 471 unique SKU demand time series were generated. These series were normalized using the Min-Max method to focus on the shape and dynamics of demand, rather than on absolute volumes. A comparative analysis of three clustering algorithms was conducted: Ward's Agglomerative Hierarchical Clustering (AHC), k-means, and Self-Organizing Maps (SOM). The optimal number of clusters, which is 10, was determined by the elbow method. The AHC-Ward method demonstrated the best quality of partitioning, achieving the highest silhouette score with a value of 0.39. The performed grouping allowed to identify 10 sets of positions with similar dynamics. Based on the analysis of the 5 largest clusters, logically justified warehouse zoning principles were proposed, which provide for compact placement of goods with similar demand in adjacent aisles. It is shown that clustering of demand time series is an effective basis for further adaptive warehouse zoning.

Keywords: storage location assignment; time series clustering; demand dynamics; warehouse logistics; order picking; product placement; e-commerce; simulation; data processing; algorithm; forecasting.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Pang, K.-W., & Chan, H.-L. (2016). Data mining-based algorithm for storage location assignment in a randomised warehouse. *International Journal of Production Research*, 55(14), 4035–4052. <https://doi.org/10.1080/00207543.2016.1244615>
2. Kalkha, H., et al. (2024). Enhancing warehouse efficiency with time series clustering: A hybrid storage location assignment strategy. *IEEE Access*, 1. <https://doi.org/10.1109/access.2024.3386887>



3. Kalafat, İ., et al. (2021). Workload forecasting of warehouse stations: Comparison between classical time series methods and XGBoost. *Data Science and Applications*, 4(2), 19–24.
4. Wang, C., & Wang, J. (2025). Research on e-commerce inventory sales forecasting model based on ARIMA and LSTM algorithm. *Mathematics*, 13(11), 1838. <https://doi.org/10.3390/math13111838>
5. Eichenseer, P., Hans, L., & Winkler, H. (2025). A data-driven machine learning model for forecasting delivery positions in logistics for workforce planning. *Supply Chain Analytics*, 9, 100099. <https://doi.org/10.1016/j.sca.2024.100099>
6. Balvak, A., & Lashchevska, N. (2025). Intelligent approaches for enhancing warehouse efficiency: Item placement, order picking, and robotic automation. *Cybersecurity: Education, Science, Technique*, 1(29), 161–177. <https://doi.org/10.28925/2663-4023.2025.29.869>
7. Islam, M. S., & Uddin, M. K. (2023). Correlated storage assignment approach in warehouses: A systematic literature review. *Journal of Industrial Engineering and Management*, 16(2), 294. <https://doi.org/10.3926/jiem.4850>
8. Park, J., Park, C., & Hong, S. (2023). Gaussian process-based storage location assignments with risk assessments for progressive zone picking systems. *Computers & Industrial Engineering*, 109700. <https://doi.org/10.1016/j.cie.2023.109700>
9. Keung, K. L., Lee, C. K. M., & Ji, P. (2021). Data-driven order correlation pattern and storage location assignment in robotic mobile fulfillment and process automation system. *Advanced Engineering Informatics*, 50, 101369. <https://doi.org/10.1016/j.aei.2021.101369>
10. Xu, C., Zhao, M., & Li, H. (2024). Data-driven simulation methodology for exploring optimal storage location assignment scheme in warehouses. *Computers & Industrial Engineering*, 110627. <https://doi.org/10.1016/j.cie.2024.110627>
11. Kaggle. (n.d.). *eCommerce events history in electronics store*. <https://www.kaggle.com/datasets/mkechinov/ecommerce-events-history-in-electronics-store>
12. Bollu, S. S. (2024). *Anomaly detection of user behavioral events in e-commerce electronics stores using SVMs* (Bachelor's thesis). Blekinge Institute of Technology.
13. Van Heesch, R. J. A. W. M. (2022). *Predicting a customer's e-commerce purchase: A study that combines sequential pattern mining with machine learning* (Master's thesis). Tilburg University.
14. Zarinchang, A., et al. (2023). Adaptive warehouse storage location assignment with considerations to order-picking efficiency and worker safety. *Journal of Industrial and Production Engineering*, 1–20. <https://doi.org/10.1080/21681015.2023.2263009>
15. Zhang, R.-Q., Wang, M., & Pan, X. (2019). New model of the storage location assignment problem considering demand correlation pattern. *Computers & Industrial Engineering*, 129, 210–219. <https://doi.org/10.1016/j.cie.2019.01.027>
16. Abramov, V., Astafieva, M., Boiko, M., Bodnenko, D., Bushma, A., Vember, V., Hlushak, O., Zhyltsov, O., Ilich, L., Kobets, N., Kovaliuk, T., Kuchakovska, H., Lytvyn, O., Lytvyn, P., Mashkina, I., Morze, N., Nosenko, T., Proshkin, V., Radchenko, S., & Yaskevych, V. (2021). Theoretical and practical aspects of the use of mathematical methods and information technology in education and science. <https://doi.org/10.28925/9720213284km>

