СЕКЦІЯ 11
МАТЕМАТИЧНІ МЕТОДИ В ЕКОНОМІЦІ

УДК 330.46:332.72:004.6

https://doi.org/10.68134/ah.2026-11-20.028

Петренко Олександр Миколайович
аспірант кафедри фінансових технологій і підприємництва, Сумський державний університет
ORCID: 0009-0008-8491-0867

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ КЛІКСТРІМ-ЛОГІВ ДЛЯ ПАРАМЕТРИЗАЦІЇ АГЕНТ-ОРІЄНТОВАНИХ МОДЕЛЕЙ ЦИФРОВИХ ПЛАТФОРМ НЕРУХОМОСТІ

Анотація. Головним методологічним викликом під час розробки сучасних агент-орієнтованих моделей (ABM) електронних ринків є проблема емпіричного калібрування мікроповедінки учасників. Більшість існуючих симуляцій ринку нерухомості спирається на довільні припущення або синтетичні розподіли параметрів, що суттєво знижує їхню прогностичну здатність в умовах цифрової економіки. У дослідженні запропоновано прикладний підхід до побудови поведінкової логіки агентів цифрової платформи нерухомості на основі інтелектуального аналізу серверних журналів подій та потоків кліків (clickstream data mining). Розроблено конвеєр трансформації неструктурованих телеметричних логів маркетплейсу (показів у видачі, глибини пагінації, застосування фільтрів, тривалості сесій та мікроконверсій) у параметри імовірнісних дерев рішень агентів. Застосування методів кластеризації сесій дозволяє виокремити гетерогенні поведінкові архетипи користувачів, оцінити порогові значення когнітивної втоми та оцифрувати вплив позиційного зміщення (position bias) на видимість об'єктів. Запропонована процедура емпіричної ініціалізації забезпечує достовірне відтворення трансакційної воронки «знизу вгору» та формує надійний аналітичний інструментарій для сценарного тестування алгоритмів ранжування, мінімізації цифрових терть пошуку й оптимізації ліквідності сучасних PropTech-маркетплейсів.

Ключові слова: інтелектуальний аналіз даних, клікстрім, агент-орієнтоване моделювання, цифрова платформа нерухомості, PropTech, поведінкові евристики, емпіричне калібрування.

Abstract. The main methodological challenge in developing modern agent-based models (ABM) of electronic markets is the problem of empirical calibration of participants' micro-behaviour. Most existing real estate market simulations rely on arbitrary assumptions or synthetic parameter distributions, which significantly reduces their predictive capacity in the digital economy. This study proposes an applied approach to constructing the behavioural logic of agents in a digital real estate platform based on intelligent analysis of server event logs and clickstream data mining. A pipeline for transforming unstructured marketplace telemetry logs (search impressions, pagination depth, filter application, session duration and micro-conversions) into parameters of probabilistic agent decision trees has been developed. The application of session clustering methods enables the identification of heterogeneous behavioural archetypes of users, estimation of cognitive fatigue thresholds, and quantification of position bias effects on object visibility. The proposed empirical initialisation procedure ensures reliable bottom-up reproduction of the transactional funnel and provides a robust analytical toolkit for scenario testing of ranking algorithms, minimisation of digital search frictions, and liquidity optimisation of modern PropTech marketplaces.

Keywords: data mining, clickstream, agent-based modeling, digital real estate platform, PropTech, behavioral heuristics, empirical calibration.

Вступ

Перехід операцій з нерухомістю в цифровий простір перетворив онлайн-платформи (PropTech-маркетплейси) на базову інфраструктуру для зведення покупців, орендарів і

продавців [1]. У межах парадигми агент-орієнтованого моделювання (Agent-Based Modeling, ABM) подібні платформи виступають як складні адаптивні системи, де макроекономічні показники — швидкість закриття угод (Time-on-Market), конверсія та цінова динаміка — формуються виключно «знизу вгору» внаслідок локальних рішень тисяч автономних агентів за умов обмеженої раціональності [2, 3].

Якщо загальна концептуальна архітектура та інституційні правила моделі цифрової платформи можуть бути формалізовані шляхом поєднання протоколів MAIA, ODD+D, FREIDA та TRACE [4, 5], то найгострішою прикладною проблемою залишається емпірична параметризація когнітивного ядра. Традиційна ініціалізація атрибутів агентів за допомогою рівномірних або нормальних випадкових розподілів перетворює симуляцію на абстрактну теоретичну модель. Натомість сучасні вебплатформи щосекунди генерують масиви цифрових слідів — клікстрім-дані (clickstream logs), які є прямим емпіричним відбитком процесу прийняття рішень людиною перед екраном пристрою [6, 7]. В умовах високої волатильності попиту та зростання конкуренції між цифровими екосистемами інтеграція реальних даних клікстріму дозволяє перетворити ABM на високоточний data-driven симулятор для попереднього тестування продуктових і монетизаційних гіпотез без ризиків для реального трафіку платформи.

Мета і завдання

Метою дослідження є розроблення методики інтелектуального аналізу клікстрім-логів (Data Mining) для емпіричної параметризації поведінкових евристик та імовірнісних дерев рішень в агент-орієнтованій моделі цифрової платформи нерухомості.

Методи

Дослідження спирається на багатофазний підхід до емпіричної типізації агентів за Smajgl et al. [8], методи інтелектуального аналізу вебданих (Web Usage Mining / Clickstream Analytics), алгоритми неконтрольованого машинного навчання (кластеризація користувацьких сесій) та протокол верифікації на основі патернів (Pattern-Oriented Modeling, POM) [9].

Для переходу від «сирих» серверних записів (JSON/CSV логів подій фронтенду та бекенду) до числових параметрів агентів у середовищі симуляції розроблено чотириетапний конвеєр обробки даних:

Сесіонізація та реконструкція треків (Sessionization): групування атомарних HTTP-запитів і подій трекінгу за ідентифікатором користувача (user_id, session_id) у часові вікна активності (з порогом неактивності $\Delta t = 30$ хв) та побудова послідовних графів переходів між станами інтерфейсу.

Інженерія поведінкових ознак (Feature Engineering): агрегування низькорівневих подій у вектор характеристик сесії: глибина скролінгу сторінки результатів пошуку (SRP), частота зміни фільтрів за ціною/локацією, час затримки на картці оголошення (dwell time), коефіцієнт клікабельності залежно від рангу видачі (CTR by rank).

Кластеризація поведінкових профілів: застосування алгоритмів кластеризації (зокрема, k-means та DBSCAN для відсікання пошукових ботів і парсерів) у просторі виділених ознак для отримання дискретних архетипів агентів-покупців із емпіричними вагами їхньої частки в популяції.

Оцінка умовних імовірностей переходів: побудова емпіричних функцій розподілу (ECDF) та матриць переходів ланцюгів Маркова для параметризації вузлів дерев рішень агентів.

Результати та обговорення

Основним результатом застосування запропонованої методики є пряме відображення (mapping) телеметричних метрик клікстріму у параметри алгоритмічних правил поведінки агентів платформи. Схему трансформації сирих логів у обчислювальні змінні ABM систематизовано в таблиці 1.

Таблиця 1

Матриця трансформації сирих подій клікстріму у параметри поведінкової логіки ABM

Подія / метрика у логах Клікстріму (Raw Logs)	**Метод обробки (Data Mining)**	**Параметр / евристика в алгоритмічному ядрі ABM**
search_filter_apply (застосування фільтрів пошуку: параметри бюджету, площі, району)	Частотний аналіз, сумісні емпіричні розподіли	Ініціалізація векторів обмежень агента у Правилі пошуку та фільтрації на старті сесії (Ci)
srp_impression (показ у результатах пошуку), pagination_click (перехід між сторінками видачі), scroll_depth (глибина прокручування)	Аналіз виживаності (Survival Analysis) глибини перегляду	Поріг когнітивної уваги Nmax у Правилі обмеженої уваги (відсікання об'єктів за межами видимого екрана)
listing_card_click (клік на картку оголошення), позиція об'єкта у видачі (r)	Логістична регресія / крива затухання уваги (Position Bias)	Базова імовірність мікроконверсії P(Click | r) та вага візуальних бейджів («ТОП», «Перевірено»)
dwell_time (тривалість перегляду) на сторінці об'єкта, photo_gallery_swipe (гортання галереї фотографій)	Аналіз порогових квантилів залученості	Оцінка відповідності об'єкта внутрішньому порогу прийнятності (Umin) у Правилі задоволення (Satisficing)
contact_reveal (відкриття контактних даних), chat_send (надсилання повідомлення в чат), favorite_add (додавання до обраного)	Побудова матриці переходів трансакційної воронки	Імовірність генерації цільового ліда P(Lead | View) та зупинки пошукового циклу агента
Динаміка views_count (кількість переглядів) та ctr (показник клікабельності) у кабінеті продавця за період ΔT	Аналіз часових рядів та точок зміни ціни (price_drop_event — подія зниження ціни)	Чутливість продавця до дефіциту уваги у Правилі цінової адаптації (дисконт або купівля Boost-просування)

Кластеризація масивів сесій за описаним конвеєром дозволяє відмовитися від гіпотези однорідного покупця і виділити у моделі три стійкі поведінкові кластери агентів-шукачів:

«Фокусовані прагматики» (High-Intent Seekers): задають жорсткі параметри фільтрації, переглядають лише перші 10–15 позицій видачі (висока чутливість до алгоритму ранжування), характеризуються тривалим часом вивчення картки об'єкта та високою конверсією у відкриття контактів.

«Широкі дослідники ринку» (Window Shoppers / Explorers): генерують велику глибину пагінації з частими скиданнями фільтрів, проте мають вкрай низьку ймовірність генерації ліда за одну сесію, створюючи інформаційний шум для продавців (завищена кількість показів без конверсії).

«Чутливі до цінових сигналів» (Bargain Hunters): використовують сортування за зростанням ціни замість дефолтного алгоритмічного ранжування, ігнорують платні промо-мітки та переважно реагують на нещодавні знижки (price_drop).

Важливим етапом дата-майнінгу (відповідно до блоку Data evaluation фреймворку TRACE) є очищення історичних логів від ендогенного алгоритмічного зміщення платформи (algorithmic & position bias). Оскільки оголошення на першій сторінці отримують кліки не лише через свою об'єктивну привабливість, а й через сам факт перебування в топ-видачі, у моделі розмежовано ймовірність потрапляння в зону уваги P(Seen | Rank) та умовну привабливість об'єкта P(Click | Seen, Xj), де Xj — вектор локальних характеристик оголошення

відносно сусідніх карток на екрані. Така декомпозиція дозволяє проводити у симуляторі контрфактичні експерименти: змінювати логіку сортування або правила платного просування без втрати поведінкової достовірності агентів.

Висновки

Перенесення акценту на інтелектуальний аналіз клікстрім-логів вирішує фундаментальну проблему емпіричного обґрунтування агент-орієнтованих моделей цифрових платформ. Запропонований конвеєр перетворення сирих телеметричних подій на параметри дискретних евристик (глибини уваги, локального порівняння, задоволення та адаптації продавця) забезпечує безшовний зв'язок між реальною поведінкою користувачів PropTech-маркетплейсів та імітаційним моделюванням. На макрорівні це дозволяє верифікувати модель за реальними показниками конверсійної воронки та експозиції об’єктів (POM-валідація), а у прикладному бізнес-аспекті — використовувати симулятор як віртуальну лабораторію (in silico A/B testing) для проєктування алгоритмів ранжування та маршрутизації лідів, які мінімізують цифрові тертя пошуку та максимізують загальну ліквідність ринку нерухомості.

Список використаних джерел

1. Baum A., Saull A., Braesemann F. PropTech 2020: The future of real estate. Saïd Business School, University of Oxford. 2020. DOI: 10.2139/ssrn.3601242
2. Tesfatsion L. Agent-based computational economics: Growing economies from the bottom up. Artificial Life. 2002. Vol. 8, No. 1. P. 55–82. DOI: 10.1162/106454602753694765
3. Simon H. A. A behavioral model of rational choice. The Quarterly Journal of Economics. 1955. Vol. 69, No. 1. P. 99–118. DOI: 10.2307/1884852
4. Ghorbani A., Bots P., Dignum V., Dijkema G. MAIA: A framework for developing agent-based social simulations. Journal of Artificial Societies and Social Simulation. 2013. Vol. 16, No. 2. P. 9. DOI: 10.18564/jasss.2166
5. Müller B., Bohn F., Dreßler G., Groeneveld J., Klassert C., Martin R., Schlüter M., Schulze J., Weise H., Schwarz N. Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. Environmental Modelling & Software. 2013. Vol. 48. P. 37–48. DOI: 10.1016/j.envsoft.2013.06.003
6. Fradkin, A. (2015). Search frictions and the design of online marketplaces. In Auctions, market mechanisms and their applications. Springer. DOI: 10.4108/EAI.8-8-2015.2260850
7. Loberto M., Luciani A., Pangallo M. The potential of big housing data: An application to the Italian real estate market. Bank of Italy. 2018. DOI: 10.2139/ssrn.3259838
8. Smajgl A., Brown D. G., Valbuena D., Huigen M. G. A. Empirical characterisation of agent behaviours in socio-ecological systems. Environmental Modelling & Software. 2011. Vol. 26, No. 7. P. 837–844. DOI: 10.1016/j.envsoft.2011.02.011
9. Grimm V., Revilla E., Berger U., Jeltsch F., Mooij W. M., Railsback S. F., Thulke H. H., Weiner J., Wiegand T., DeAngelis D. L. Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. Science. 2005. Vol. 310, No. 5750. P. 987–991. DOI: 10.1126/science.1116681