

**О. Ю. ЧЕРЕДНІЧЕНКО, О. В. ІВАЩЕНКО, Ю. М. ГОНТАР, Б. М. ВОРОНА**

### **ИНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ПРОПОЗИЦІЙ ТОВАРІВ НА ОСНОВІ КОНТЕКСТНИХ РЕКОМЕНДАЦІЙ**

Інтернет-технології є невід'ємною складовою відносин, які виникають у сучасному суспільстві. Через швидке впровадження та зручність електронних майданчиків, прогнозовано зростає попит на ринку ІТ-продуктів для рекомендаційних систем. У статті розглянуті різноманітні обмеження поточних рекомендаційних методів та обговорено можливі розширення, що можуть покращити рекомендаційні можливості та зробити їх більш цінними для широкого кола додатків. Ці розширення включають покращення сприймання користувачів та елементів, включення контекстної інформації в рекомендаційний процес, підтримка багатокритеріальних рейтингів та надання більш гнучких і водночас менш нав'язливих типів рекомендацій. Важливу роль відіграє інтеграція діяльності, яка полягає у підтримці усіх аспектів електронної комерції від виконання транзакцій до підтримки мережі постачання, що дає змогу спростити документообіг та збільшити вигоду учасників. Направленість даної розробки – проводити аналітичну обробку даних торговельних майданчиків, на основі контекстних рекомендацій, об'єктивний аналіз та здійснювати актуальний моніторинг ділової активності на торговельному майданчику. Розглянуто задачу складання різноманітних аналітичних звітів, що дозволить учасникам ринку ІТ-продуктів для рекомендаційних систем об'єктивно і своєчасно аналізувати розвиток ситуації на ринку, виявляти існуючі та прогнозні тенденції. Побудова сфери надання інтелектуальних аналітичних послуг здійснюється для залучення додаткових учасників, або якісно нових гравців ринку та одержання додаткового прибутку. Для обробки доцільно використовувати принципово нові технології Data Mining, що дозволить отримати якісно цінні дані. Data Mining – це технологія, призначена для пошуку у великих інформаційних масивах неочевидних даних, об'єктивних, корисних на практиці закономірностей.

**Ключові слова:** електронні торговельні майданчики, рекомендаційні системи, контекстна інформація, інтерпретація даних, моделі, програмний компонент, інтелектуальний аналіз даних.

**О. Ю. ЧЕРЕДНІЧЕНКО, О. В. ІВАЩЕНКО, Ю. М. ГОНТАР, Б. М. ВОРОНА**

### **ИНТЕЛЕКТУАЛЬНЫЙ АНАЛИЗ ПРЕДЛОЖЕНИЙ ТОВАРОВ НА ОСНОВЕ КОНТЕКСТНЫХ РЕКОМЕНДАЦИЙ**

Интернет-технологии являются неотъемлемой составляющей отношений, возникающих в современном обществе. Быстрое внедрение и удобство электронных площадок спровоцировало прогнозируемый рост спроса на рынке IT-продуктов для рекомендательных систем. В статье рассмотрены различные ограничения текущих рекомендательных методов и обсуждены возможные расширения, которые могут улучшить рекомендательные возможности и сделать их более ценными для широкого круга приложений. Эти расширения включают улучшение восприятия пользователей и элементов, включение контекстной информации в рекомендательный процесс, поддержка многокритериальных рейтингов и предоставления более гибких и одновременно менее навязчивых типов рекомендаций. Важную роль играет интеграция деятельности, которая заключается в поддержке всех аспектов электронной коммерции от выполнения транзакций до поддержки сети поставок, что позволяет упростить документооборот и увеличить выгоду участников. Направленность данной разработки – проводить аналитическую обработку данных торговых площадок, на основе контекстных рекомендаций, объективный анализ и осуществлять актуальный мониторинг деловой активности на торговой площадке. Рассмотрена задача составления разнообразных аналитических отчетов, что позволит участникам рынка IT-продуктов для рекомендательных систем объективно и своевременно анализировать развитие ситуации на рынке, выявлять существующие и прогнозные тенденции. Построение сферы предоставления интеллектуальных аналитических услуг осуществляется для привлечения дополнительных участников, или качественно новых игроков рынка и получения дополнительной прибыли. Для обработки целесообразно использовать принципиально новые технологии Data Mining, что позволит получить качественно ценные данные. Data Mining – это технология, предназначенная для поиска в больших информационных массивах неочевидных данных, объективных, полезных на практике закономерностей.

**Ключевые слова:** электронные торговые площадки, рекомендательная система, контекстная информация, интерпретация данных, модели, программный компонент, интеллектуальный анализ данных.

**O. Y. CHEREDNICHENKO, O. V. IVASHCHENKO, Y. M. GONTAR, B. M. VORONA**

### **DATA MINING OF COMMODITY PROPOSALS BASED ON CONTEXT RECOMMENDATIONS**

Internet technologies are an integral part of the relationship that arises in modern society. The rapid introduction and convenience of electronic platforms triggered the projected growth in demand in the market for IT products for recommender systems. The article discusses various limitations of current recommender methods and discusses possible extensions that can improve the recommender capabilities and make them more valuable for a wide range of applications. These extensions include improving the perception of users and elements, including contextual information in the recommendatory process, supporting multi-criteria ratings and providing more flexible and at the same time less intrusive types of recommendations. When integrating the relevant information technology to develop a commodity proposals environment, it is therefore necessary to consider the personalization requirements of the proposal to ensure that the technology achieves its intended result. This study therefore sought to apply context aware technology and recommendation algorithms to develop a system realize personalized goals in a context aware manner and improve commodity proposals effectiveness. In order to offer context-aware and personalized information, intelligent processing techniques are necessary. Different initiatives considering many contexts have been proposed, but users preferences need to be learned to offer contextualized and personalized services, products or information. Therefore, this paper proposes an agent-based architecture for context-aware and personalized event recommendation based on ontology and the spreading algorithm. The use of ontology allows to define the domain knowledge model, while the spreading activation algorithm learns user patterns by discovering user interests. Also from the statistical observation, it is found that there exists a higher level agreement towards the system between the participants of both end users and experts.

**Keywords:** electronic market places, recommender system, context information, data interpretation, model, software component, data mining.

**Вступ.** Останнім часом ринок електронної секторі. Українські компанії все частіше купують комерції в Україні активно розвивається, причому не товари та послуги на електронних торговельних тільки в споживацькому, але й у корпоративному майданчиках (ЕТП). З впровадженням у дію Закону

України «Про здійснення державних закупівель», який оголошує обов'язкове проведення всіх державних закупівель у електронній формі, очікується, що обсяги ринку ЕТП зростуть в кілька разів. Зарано декларувати швидке та повне переміщення підприємств із паперової форми закупок до електронної, але все ж у найближчій перспективі в Україні корпоративний сектор стане так само активним в електронній комерції, як і споживачі.

Загалом, електронна комерція розвивається в чотирьох напрямках:

1. C2C (customer to customer – споживач для споживачів) системи електронної торгівлі між приватними особами (наприклад, сайт-дошка оголошень Slando.ua).
2. B2C (business to customer – бізнес для споживачів) – електронні майданчики для споживачів, тобто онлайн магазини, в яких кожен з нас купує товари масового попиту, побутову техніку.
3. B2B (business to business – бізнес для бізнесу) – електронні майданчики, на яких постачальники і закупівельники можуть вести торгово-закупівельну діяльність, купуючи товари, роботи і послуги для потреб компанії, починаючи від оргтехніки і закінчуючи важким металургійним обладнанням.
4. B2G (business to government – бізнес для держави) – електронні майданчики, на яких державні установи можуть здійснювати закупівлі відповідно до законодавства про держзакупівлі, купуючи товари, роботи і послуги для потреб державного сектору.

На ринку України представлені такі види електронних майданчиків:

- Незалежні електронні торгові майданчики. Пропонують проводити електронні торги на їх платформі, надаючи певний функціонал.
- Внутрішньо корпоративні електронні торгові майданчики. Деякі компанії організовують власні електронні майданчики (як приклад можна привести торговий майданчик компанії ДТЕК – tenders.dtek.com).

Корпоративні майданчики зручні для роботи конкретного підприємства, оскільки враховують всю специфіку його закупівельної діяльності. Що ж стосується незалежних ЕТМ, то завдяки запропонованій індивідуальній системі налаштувань, практично будь-яке підприємство може працювати на незалежній ЕТМ, не обмежуючи себе в функціоналі.

Переваги, які надаються користувачам системами електронної торгівлі:

1. Електронні торгові майданчики зазвичай надають відкритий доступ до баз даних організацій і торгових процедур. Це дає можливість швидко знайти необхідну інформацію про контрагентів, конкурентів, торгові процедури, ціни і т. п., а також швидко і репрезентативно провести маркетингові дослідження.

2. Реєстрація в системі означає розміщення інформації про вашу організацію (реквізити, контакти, сайт), про пропоновану вами продукцію і (або) послуги (з малюнками, фотографіями та ін.). Це дозволяє без особливих витрат проводити масштабну рекламну кампанію по їх просуванню.
3. Зазвичай надається можливість підписатися на профільну розсилку і отримувати по електронній пошті повідомлення про цікаві процедури.
4. Організація може брати участь в різних процедурах будь-якого рівня складності (конкурентні переговори, запити цін / пропозицій, аукціони покупця / продавця, різні види конкурсів).
5. Зазвичай надається безкоштовний доступ до документації в електронному вигляді по всіх торгових процедурах.
6. Скорочення витрат на участь в торгових процедурах.
7. Мінімізація паперового документообігу (документи формуються в електронному вигляді і підписуються електронним цифровим підписом).
8. Значне розширення ринків збуту (зняття географічних бар'єрів) веде до значного збільшення продажів.
9. Системи надають зручні сервіси для ефективної роботи (отримання електронного цифрового підпису, банківської гарантії, експрес-кредитування на внесення забезпечення заявок, послуги страхування, розрахунок логістики, переклад на іноземні мови та ін.).

Через швидке впровадження та зручність електронних майданчиків, прогнозовано зростає попит на ринку ІТ-продуктів для рекомендаційних систем, які стали важливою областю досліджень з моменту появи перших документів про спільну фільтрацію в середині 1990-х років [1]. Інтерес до цієї області все ще залишається високим, оскільки вона являє собою багатопроблемну область досліджень і має велику кількість практичних застосувань, які допомагають користувачам справлятися з перевантаженням інформацією та надавати персоналізовані рекомендації, зміст та послуги для них. Приклади таких програм включають рекомендації книг, компакт-дисків та інших продуктів на Amazon.com, фільмів MovieLens та новин VERSIFI Technologies (раніше AdaptiveInfo.com) [2]. Крім того, деякі постачальники включили можливість рекомендацій у свої комерційні сервери [3].

Однак, незважаючи на всі ці досягнення, існуюче покоління рекомендаційних систем все ще потребує додаткових покращень, щоб зробити методи рекомендацій більш ефективними та застосовуваними до ще більш широкого спектру додатків у реальному житті, включаючи рекомендації щодо відпусток, певних видів фінансових послуг для інвесторів, та продуктів для покупки в магазині, які створені "розумним" кошиком [4]. Ці вдосконалення включають в себе кращі методи представлення поведінки користувачів

та інформації про рекомендовані елементи, більш просунуті методи моделювання рекомендацій, включення різноманітної контекстної інформації в процес рекомендації, використання багатокритеріальних рейтингів, розробка менш нав'язливих та більш гнучких методів рекомендацій, які також покладаються на заходи, які більш ефективно визначають ефективність рекомендаційних систем.

Об'єктом даного дослідження є процес розширення можливостей рекомендаційних систем. Рекомендаційні системи можуть бути розширені кількома способами, які включають в себе поліпшення розуміння користувачів і елементів, включення контекстної інформації в процес рекомендації, підтримку оцінок з множинними критеріями і надання більш гнучких і менш нав'язливих рекомендацій. Такі більш всеосяжні моделі рекомендаційних систем можуть забезпечити найкращі рекомендації.

**Дослідження існуючих рішень проблем.** За останні кілька років було проведено велику кількість досліджень на тему рекомендаційних технологій, які використовували широкий спектр статистичних методів, машинного навчання, пошуку інформації та інших, які значно просунули сучасний рівень в порівнянні з ранніми рекомендаційними системами, що використовують спільну та контентно-орієнтовану евристику. Рекомендаційні системи можуть бути класифіковані наступним чином:

- контентно-орієнтовані, колаборативні, гібридні, засновані на використовуваних рекомендаційних підходах
- евристичні або модельні, засновані на основних типах методів рекомендації, які використовуються для оцінки рейтингу.

Ці два ортогональних аспекти для класифікації рекомендаційних систем представлено в табл. 1.

Проте як колаборативні, так і контентно-орієнтовані методи мають певні обмеження. Більш того, для того, щоб забезпечити кращі рекомендації та мати можливість використовувати рекомендаційні системи в набагато більш складних типах програм, таких як рекомендації щодо відпусток чи певних видів фінансових послуг, більшість методів, розглянутих у цьому розділі, потребують значного розширення. Наприклад, навіть для традиційної програми рекомендацій фільму [3] показано, що розширивши традиційний підхід до спільного використання фільтрів на основі пам'яті, враховуючи контекстну інформацію, таку як, коли, де і з ким було побачено фільм, результати рекомендаційної системи можуть перевершити чистий традиційний метод колаборативного фільтрування. Багато реалістичних рекомендаційних додатків, в тому числі декілька ділових додатків, є, напевно, більш складними, ніж система, що рекомендує фільм, і вимагають урахування інших факторів у розгляді рекомендацій. Тому необхідність розробки більш просунутих методів рекомендації ще більш актуальна для таких типів додатків.

#### **Всебічне розуміння користувачів та елементів.**

Як зазначалося в [2], [3], [4], більшість рекомендаційних методів дають оцінки, засновані на обмеженому

розумінні користувачів і елементів, які враховуються в профілях користувачів і елементів, і не приймають повну перевагу інформації в транзакційних історіях користувача та інших доступних даних. Незважаючи на певний прогрес, досягнутий при впровадженні профілів користувачів і позицій в деякі з більш ранніх методів рекомендацій, ці профілі все ще мають тенденцію бути досить простими і не використовують деякі з більш просунутих методів профілювання. На додаток до використання традиційних функцій профілю, таких як ключові слова і прості демографічні дані користувача, більш просунуті методи профілювання, засновані на правилах інтелектуального аналізу даних, послідовності і підписів які описують інтереси користувача, можуть використовуватися для створення профілів користувачів. Крім того, на додаток до використання традиційних функцій профілю елементів, таких як ключові слова, аналогічні передові методи профілювання також можуть використовуватися для створення всеосяжних профілів елементів. Що стосується рекомендаційних систем, то сучасні методи профілювання, засновані на інтелектуальному аналізі даних, використовуються головним чином в контексті аналізу використання Інтернет, тобто для виявлення навігаційних шаблонів використання Інтернету (послідовності перегляду сторінок) користувачами, щоб забезпечити найкращі рекомендації веб-сайтів; однак такі методи не набули широкого поширення в рейтингових системах [5].

Після створення профілів користувачів і елементів найбільш загальна функція оцінки рейтингів може бути визначена в термінах цих профілів і раніше заданих рейтингів наступним чином:

Нехай профіль користувача визначається як вектор функції  $p$ , тобто  $C_i = (a_{i1}, \dots, a_{ip})$ . Крім цього, нехай профіль елемента  $j$  визначається як вектор  $r$  ознак, тобто  $S_j = (b_{j1}, \dots, b_{jr})$ . окрім цього, нехай  $C$  – вектор усіх профілів користувачів, тобто  $C = (C_1, \dots, C_m)$ , і нехай  $S$  – вектор всіх профілів елементів, тобто  $S = (S_1, \dots, S_n)$ . Тоді найбільш загальну процедуру оцінки рейтингу можливо визначити так:

$$r'_{ij} = \begin{cases} r_{ij}, & \text{if } r_{ij} \neq \emptyset, \\ u_{ij}(R, c, s), & \text{if } r_{ij} = \emptyset. \end{cases} \quad (1)$$

Ця процедура оцінює кожну невідому оцінку  $r'_{ij} = u_{ij}(R, c, s)$  у термінах відомих оцінок  $R = \{r_{ij} \neq \emptyset\}$ , профілів користувачів  $C$  і профілів елементів  $S$ . Ми можемо використовувати гнучкі методи для оцінки функції корисності  $u_{ij}$ , включаючи різні евристики, найближчі класифікатори, дерева рішень, гнучкі методи, радіальні базові функції, регресії, нейронні мережі і методи реляційного навчання (оскільки  $C$  і  $S$  можуть бути матрицями або реляційними таблицями) [6]. Більш того, ми хотіли б відзначити, що (1) представляє найбільш загальну модель, яка залежить від цілого діапазону входів, включаючи характеристики користувача ( $C_i$ ) і можливих інших користувачів  $C = (C_1, \dots, C_m)$ , характеристики  $S_j$  елемента  $j$  й, можливо,

Таблиця 1 – Класифікація рекомендаційних систем

Рекомендаційний підхід	Рекомендаційна техніка	
	Евристична основа	На основі моделі, яка зазвичай використовується
На основі вмісту	<p>Методи, що зазвичай використовуються:</p> <ul style="list-style-type: none"> <li>• TF-IDF (пошук інформації)</li> <li>• Кластеризація</li> </ul> <p>Репрезентація дослідницьких прикладів:</p> <ul style="list-style-type: none"> <li>• Lang, 1995</li> <li>• Balabanovic &amp; Shoham, 1997</li> <li>• Pazzani &amp; Billsus, 1997</li> </ul>	<p>Баєсівські класифікатори</p> <ul style="list-style-type: none"> <li>• Кластеризація</li> <li>• Рішення дерев</li> <li>• Штучні нейронні мережі</li> </ul> <p>Приклади репрезентативних досліджень:</p> <ul style="list-style-type: none"> <li>• Parzani &amp; Billsus, 1997</li> <li>• Mooney et al., 1998</li> <li>• Mooney &amp; Roy, 1999</li> <li>• Billsus &amp; Pavan, 1999</li> </ul>
Спільна робота	<p>Методи, що зазвичай використовуються:</p> <ul style="list-style-type: none"> <li>• Найближчий сусід (косинус кореляція)</li> <li>• Кластеризація</li> <li>• Теорія графів</li> </ul> <p>Приклади репрезентативних досліджень:</p> <ul style="list-style-type: none"> <li>• Resnick et al., 1994</li> <li>• Hill et al., 1995</li> <li>• Shardanand &amp; Macs, 1995</li> <li>• Breese et al., 1998</li> <li>• Nakamura &amp; Abc, 1998</li> <li>• Aggarwale et al., 1999</li> <li>• Delgado &amp; Ishii, 1999</li> <li>• Pennock &amp; Horwitz, 1999</li> <li>• Sarwar et al., 2001</li> </ul>	<p>Методи, що зазвичай використовуються:</p> <ul style="list-style-type: none"> <li>• Найближчий сусід (косинус кореляція)</li> <li>• Кластеризація</li> <li>• Теорія графів</li> </ul> <p>Приклади репрезентативних досліджень:</p> <ul style="list-style-type: none"> <li>• Billsus &amp; Pazzani, 1998</li> <li>• Breese et al., 1998</li> <li>• Ungar &amp; Foster, 1998</li> <li>• Chien &amp; George, 1999</li> <li>• Getoor &amp; Sahami, 1999</li> <li>• Pennock &amp; Horwitz, 1999</li> <li>• Goldberger et al., 2001</li> <li>• Kumaret al., 2001</li> <li>• Pavlov &amp; Pennock, 2002</li> <li>• Shaniet al., 2002</li> <li>• Vuet al., 2002, 2004</li> <li>• Hofmann 2003, 2004</li> <li>• Marlin, 2003</li> </ul>
Гібридна модель	<p>Поєднання компонентів на основі вмісту та спільних компонентів за допомогою:</p> <ul style="list-style-type: none"> <li>• Лівійна комбінація прогнозованих оцінок</li> <li>• Різні схеми голосування</li> <li>• Включення одного компоненту у складі</li> </ul> <p>Евристичні методи для інших дослідницьких прикладів дослідження:</p> <ul style="list-style-type: none"> <li>• Balabanovic &amp; Shoham, 1997</li> <li>• Claypoole et al., 1999</li> <li>• Good et al., 1999</li> <li>• Pazzani, 1999</li> <li>• Billsus &amp; Pazzani, 2000</li> <li>• Iran &amp; Cohen, 2000</li> <li>• Melville et al., 2002</li> </ul>	<p>Поєднання компонентів на основі вмісту та спільної роботи:</p> <ul style="list-style-type: none"> <li>• Включення одного компоненту у вигляді панорами моделі для іншого</li> <li>• Побудова однієї об'єднуючої моделі</li> </ul> <p>Репрезентативні дослідження:</p> <ul style="list-style-type: none"> <li>• Basualal, 1998</li> <li>• Condliff et al., 1999</li> <li>• Soboroff &amp; Nicholas, 1999</li> <li>• Ansaric et al., 2000</li> <li>• Popesculet al., 2001</li> <li>• Scheinet al., 2002</li> </ul>

інші елементи  $S = (S_1, \dots, S_n)$ , рейтинги (переваги)  $R_i$ , виражені користувачем  $i$ , і рейтинги (переваги), виражені усіма іншими користувачами  $R = \{r_{ij} \neq \emptyset\}$ . Таким чином, функція  $u_{ij}$  явно включає в себе спільні, засновані на контенті і гібридні методи. Однак більшість існуючих рекомендаційних систем роблять функцію  $u_{ij}$  залежною тільки від малої підмножини всього вхідного простору  $R, c, s$ . Наприклад, функція  $u_{ij}$  для традиційних методів спільної фільтрації на основі пам'яті не залежить від входів  $C$  і  $S$ , і обмежує  $R$  тільки стовпчиком  $R_j$  і зазвичай тільки для множини  $N$  найближчих сусідів  $r_{ij}$  для стовпця  $R_j^2$ .

Цікавим завданням дослідження було б розширити профілі на основі атрибутів, як визначено  $C$  і  $S$ ,

для використання більш складних методів профілювання, описаних вище, таких як правила, послідовність і методи на основі сигнатур.

**Розширення для рекомендаційних технік на основі моделі.** Деякі з підходів, заснованих на моделях, забезпечують суворі методи рейтингової оцінки з використанням різних методів статистичного і машинного навчання. Однак інші області математики та інформатики, такі як теорія математичного наближення [8], можуть також сприяти розробці кращих рейтингових методів оцінки, які визначаються формулою (1). Одним із прикладів підходу на основі апроксимації до визначальної функції  $u_{ij}$  в (1) є радіальні базисні функції, які визначаються наступним чином: дано множину точок  $X = \{x_i, \dots, x_m\}$  (де  $x_i \in IR^N$ ) і

значення невідомої функції  $f$  (наприклад, рейтингової функції) в цих точках, тобто  $f(x_1), \dots, f(x_m)$  радіальна базисна функція  $r_{f,x}$  оцінює значення  $f$  в усьому  $IR^N$ , враховуючи  $r_{f,x}(x_i) = f(x_i)$  для всіх  $i = 1, \dots, m$ , як

$$r_{f,x}(x) = \sum_{i=1}^m \alpha_i \varphi(\|x - x_i\|), \quad (2)$$

де функція, яка задовольняє вимогам

$$\sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j \varphi(\|x_i - x_j\|) > 0 \quad (3)$$

для всіх різних точок в та всіх коефіцієнтів.

Тоді відома теорема [6] стверджує, що якщо  $\varphi$  є додатно визначеною функцією, то існує єдина функція виду (2), що задовольняє умовам для всіх. Деякі популярні приклади функцій:

1.  $\varphi(r) = r^\beta$ , де  $\beta > 0$  додатне непарне число;
2.  $\varphi(r) = r^k \log(r)$ , де  $k \in IN$  (тонка сплайн-поверхня);
3.  $\varphi(r) = e^{1-\alpha r^2}$ , де  $\alpha > 0$  (Гаус).

Однією з переваг радіальних базисних функцій є те, що вони широко вивчені в теорії наближень, а їх теоретичні властивості та використання радіальних базисних функцій у багатьох практичних застосуваннях були добре зрозумілими [6], [7]. Тому було б цікаво застосувати їх до оцінки невідомих рейтингів у рекомендаційних системах.

Одне з попереджень з використанням радіальних базисних функцій у рекомендаційних системах полягає в тому, що простір рекомендацій  $C \times S$  зазвичай не представляє  $N$ -вимірний евклідовий простір. Тому однією дослідницькою задачею є розповсюдження радіальних базисних методів від реальних чисел до інших доменів та їх застосування до проблем рекомендаційних систем. Застосовність інших методів апроксимації для оцінки в (1) представляє ще одну цікаву тему дослідження.

**Багатомірність рекомендацій.** Нинішнє покоління рекомендаційних систем діє в двовимірному просторі. Тобто вони роблять свої рекомендації тільки на основі інформації про користувачів, елементи і не враховують додаткову контекстну інформацію, яка може мати вирішальне значення в деяких додатках. Однак, у багатьох ситуаціях, корисність певного продукту для користувача може значно залежати від часу (наприклад, пори року, сезону, місяцю або дня тижня). Це може також залежати від людини, з якою продукт буде споживатися або ділитися, і за яких обставин. У таких ситуаціях може бути недостатньо просто рекомендувати користувачам елементи; система рекомендацій повинна приймати до уваги при рекомендації продукту додаткову контекстну інформацію, таку як час, місце і компанія користувача. Наприклад, рекомендуючи пакет відпустки, система повинна також

враховувати пору року, яким користувач планує подорожувати, умови поїздки та обмеження на той момент і іншу контекстну інформацію.

Як інший приклад, користувач може мати істотно різні переваги у фільмах, які він хоче побачити, коли виїжджає до кінотеатру з другом в суботу ввечері, і тих, які планує дивитись вдома з батьками ввечері середі.

Як стверджувалося в [9] і [10], важливо розширити традиційні двовимірні й Користувач  $\times$  Елемент рекомендаційні методи для багатовимірних налаштувань. Крім того, включення знань про завдання користувача в алгоритм рекомендації в певних програмах може привести до кращих рекомендацій.

Щоб взяти до уваги контекстну інформацію, [2] пропонує визначити функцію корисності (або оцінок) над багатовимірним простором  $D_1 \times \dots \times D_n$  (на відміну від традиційного двовимірного) як

$$u: D_1 \times \dots \times D_n \rightarrow R. \quad (4)$$

Тоді проблема з рекомендаціями визначається вибором певних «що» вимірювань  $D_{i1} \dots D_{ik} (k < n)$  і певних «для кого» вимірювань  $D_{j1} \dots D_{jl} (l < n)$ , які не перекриваються, тобто  $\{D_{i1} \dots D_{ik}\} \cap \{D_{j1} \dots D_{jl}\} = \emptyset$ , і рекомендують, для кожного набору  $(d_{j1}, \dots, d_{jl}) \in D_{j1} \times \dots \times D_{jl}$ , набір  $(d_{i1}, \dots, d_{ik}) \in D_{i1} \times \dots \times D_{ik}$ , який максимізує корисність  $u(d_1, \dots, d_n)$ , тобто

$$\forall (d_{j1}, \dots, d_{jl}) \in D_{j1} \times \dots \times D_{jl},$$

$$\begin{aligned} (d_{i1}, \dots, d_{ik}) &= \\ &= \operatorname{argmax}_{\substack{(d'_{i1}, \dots, d'_{ik}) \in D_{i1} \times \dots \times D_{ik} \\ (d'_{j1}, \dots, d'_{jl}) = (d_{j1}, \dots, d_{jl})}} u(d'_1, \dots, d'_n). \end{aligned}$$

Наприклад, в разі системи рекомендацій для фільму потрібно враховувати не тільки характеристики фільму  $d_1$  і людини, яка хоче подивитися фільм  $d_2$ , але також і таку контекстну інформацію, як:

- 1)  $d_3$ : де і ким буде переглянуто фільм (наприклад, у кінотеатрі, вдома по телевізору, на відео або DVD),
- 2)  $d_4$ : з ким буде переглянуто фільм (наприклад, поодиноці, з подругою / другом, друзями, батьками і т. д.),
- 3)  $d_5$ : коли фільм буде переглянуто (наприклад, в будні дні або у вихідні дні, вранці / удень / увечері, в ніч відкриття і т. д.).

Як обговорювалося раніше, кожен з компонентів  $d_1, d_2, d_3, d_4, d_5$  можна визначити як вектор його характеристик і загальна корисна функція  $u(d_1, d_2, d_3, d_4, d_5)$  може бути досить складною і враховувати різні ефекти взаємодії між векторами  $d_1, d_2, d_3, d_4, d_5$ .

Як стверджувалося в [11], багато двовимірних алгоритмів рекомендацій не можуть бути безпосередньо поширені на багатовимірний випадок. Крім того, пропонується підхід, заснований на рекомендаціях щодо скорочення, який використовує тільки рейтинги, які стосуються контексту зазначених користувачем

критеріїв, в яких робиться рекомендація. Наприклад, щоб рекомендувати фільм для людини, яка хоче побачити його в кінотеатрі в суботу ввечері, підхід, заснований на скороченні, буде використовувати тільки доступні рейтинги фільмів, які можна побачити в кінотеатрах у вихідні, якщо це визначено даними, що місце і час тижня впливають на поведінку кіноглядачів. Вибравши тільки рейтинги, які стосуються контекстної рекомендації, підхід, заснований на скороченні, проектує багатовимірний куб рейтингів за двома основними параметрами користувача і елемента. Потім для отримання рекомендації можна використовувати будь-який стандартний метод двовимірних рекомендацій. Оскільки ці рекомендації засновані тільки на контекстно-залежному наборі рейтингів, це зводиться до створення локальної моделі, що створює контекстні рекомендації.

Іншим можливим підходом до розробки багатовимірних рекомендацій було б розгортання ієрархічного баєсівського методу, представленого в [3], який може бути розширений з двох до багатовимірного випадку наступним чином: замість розгляду двовимірного випадку, як визначено в (1), де до характеристики користувача  $d_1$ , які визначаються вектором  $\mathbf{z}_i$ , і характеристик елемента  $d_2$  з вектором  $\mathbf{w}_j$ , ми також можемо додати контекстні розміри  $d_3, \dots, d_n$ , де  $\mathbf{d}_i = (d_{i1}, \dots, d_{ix_i})$  – вектор характеристик розмірності  $D_i$ . Тоді функція  $r = u(d_1, d_2, \dots, d_n)$  поширюється від (1) до лінійної комбінації  $d_1, d_2, \dots, d_n$ , і також включає ефекти взаємодії між цими розмірностями (тобто ефект взаємодії, який визначається матрицею  $\{x_{ij}\}$  в (1), повинен бути розширеним). Одне із завдань дослідження – зробити ці розширення масштабованими при великих значеннях  $n$ .

**Багатокритеріальні рейтинги.** Більшість сучасних рекомендаційних систем відносяться до систем з одним рейтингом, такі як рейтинги фільмів і книг. Однак в деяких додатках, таких як рекомендації для ресторанів, край важливо включити рейтинги з декількома критеріями в рекомендації. Наприклад, багато путівники по ресторанах, такі як Zagat's Guide, надають три критерії для оцінки ресторанів: їжа, декор та обслуговування [12]. Хоча багатокритеріальні оцінки ще не були розглянуті в літературі по рекомендаційним системам, вони були широко вивчені в співтоваристві Operations Research [13]. Типові рішення проблем оптимізації багатокритеріальних задач:

1. Знайти оптимальні рішення Парето,
2. Взяти лінійну комбінацію декількох критеріїв і зменшити проблему до задачі оптимізації з одним критерієм,
3. Оптимізувати найважливіший критерій і перетворити інші критерії в обмеження;
4. Послідовна оптимізація одного критерію за раз, перетворення оптимального рішення в обмеження (обмеження) і повторення процесу для інших критеріїв.

Щоб проілюструвати, як деякі з цих методів можуть бути використані в рекомендаційних системах,

розглянемо застосування підходу 3 до проблеми рекомендації ресторанів  $r$  користувачеві  $s$  на основі критеріїв користувача з якості їжі  $f_c(r)$ , декору  $d_c(r)$ , і обслуговування  $s_c(r)$ . Ми можемо вважати якість їжі  $f_c(r)$  основним критерієм і використовувати інші в якості обмежень, тобто ми хочемо знайти ресторани  $r$ , які максимізують  $f_c(r)$ , за умови, що обмеження  $d_c(r) > \alpha_c$ , і  $s_c(r) > \beta_c$ , де  $\alpha_c$  і  $\beta_c$  – мінімальні рейтинги для декору і обслуговування (наприклад, користувач  $s$  не ходитиме в будь-який ресторан з оцінкою декору і обслуговування нижче 10, з можливих 30, незалежно від того, якої якості їжі там). Ця проблема ускладнюється тим фактом, що у нас зазвичай не буде оцінок декору  $d_c(r)$  і обслуговування  $s_c(r)$  для всіх ресторанів. Потім завдання системи рекомендацій полягає в тому, щоб оцінити невідомі оцінки  $d'_c(r)$  і  $s'_c(r)$  і, наприклад, використовуючи методи оцінки, і знайти всі ресторани  $r$ , що задовольняють обмеженням  $d'_c(r) > \alpha_c$ , і  $s'_c(r) > \beta_c$ . Як тільки ми знайдемо всі ресторани, що задовольняють обмеженням за цими оціночними рейтингами, ми можемо використовувати ці ресторани в пошуках максимуму  $f_c(r)$ . Однак, як і у випадку з оцінками відділів і сервісів, у нас можуть не бути рейтингів продуктів харчування  $f_c(r)$  для всіх таких ресторанів  $f_c(r)$  і, отже, також необхідно буде використовувати процедуру оцінки рейтингу перед тим, як робити будь-які рекомендації. Ми вважаємо, що процес знаходження оптимального набору Парето та ітеративного методу послідовних оптимізацій одного критерію для згаданих вище рекомендацій по множинним критикам також являє собою цікаві і складні проблеми.

**Невтручання (nonintrusiveness).** Багато рекомендаційних системи є нав'язливими в тому сенсі, що вони вимагають явного зворотного зв'язку від користувача і часто на значному рівні участі користувача. Наприклад, перш ніж рекомендувати будь-які статті в групах новин, системі необхідно отримати рейтинги раніше прочитаних статей. Оскільки недоцільно вимагати рейтинги всіх цих статей від користувача, деякі системи рекомендацій використовують ненав'язливі методи визначення рейтингів, коли певні проксі використовуються для оцінки реальних рейтингів. Наприклад, кількість часу, який користувач витрачає на читання статті в групі новин, може служити проксі-рейтингами цієї статті. Деякі ненав'язливі методи отримання зворотного зв'язку з користувачем представлені в [14], [15]. Проте, ненав'язливі рейтинги (наприклад, час, витрачений на читання статті) часто є неточними і не можуть повністю замінити явні оцінки, надані користувачем. Таким чином, проблема мінімізації втручання при збереженні певних рівнів точності рекомендацій повинна вирішуватися дослідниками рекомендаційних систем.

Один із способів вивчення проблеми втручання – визначити оптимальну кількість оцінок, які система повинна запитувати у нового користувача. Наприклад, перш ніж рекомендувати будь-які фільми, MovieLens.org спочатку просить користувача оцінити

зумовлену кількість фільмів (наприклад, 20). Цей запит несе певні витрати для кінцевого користувача, який може бути змодельований по-різному, причому найпростіша модель являє собою модель з фіксованою вартістю (тобто, вартість рейтингу кожного фільму дорівнює  $C$ , а вартість рейтингу  $n$  фільмів  $C_n$ ). Тоді проблему нав'язливості можна сформулювати як проблему оптимізації, яка намагається знайти оптимальну кількість оцінок запитів  $n$  наступним чином: кожен додатковий рейтинг, наданий користувачем, підвищує точність рекомендацій і, отже, призводить до певних переваг для користувача. Однією з цікавих задач, пов'язаних з проблемою втручання, було б розробка формальної моделі для визначення та вимірювання переваг  $B(n)$  для надання  $n$  початкових рейтингів з точки зору підвищення точності прогнозів на основі цих оцінок. Як тільки стає відомо, як вимірювати переваги  $B(n)$  (наприклад, вимірюючи точність прогнозування рекомендаційної системи), потрібно визначити оптимальну кількість початкових оцінок  $n$ , яка максимізує вираз  $B(n) - C(n)$ . Ясно, що оптимальне значення  $n$  досягається, коли граничні вигоди дорівнюють граничним витратам, тобто коли  $\Delta B(n) = C$ . Оптимальне рішення повинно існувати в припущенні, що  $B(n)$  є монотонно зростаючою функцією по  $n$  зі зменшуваними граничними перевагами  $\Delta B(n)$ , що асимптотично сходиться до нуля.

Ще одна цікава можливість дослідження полягає в розробці моделей граничних витрат, які є більш просунутими, ніж модель з фіксованою вартістю, описана вище, і яка потенційно може включати аналіз витрат і переваг використання як неявних, так і явних оцінок в системі рекомендацій. Нарешті, проблема поступового вибору навчальних даних для цілей моделювання – проблема активного навчання, яка є досить добре вивченою областю в літературі машинного навчання, і були запропоновані численні підходи до вирішення цієї проблеми [16], [17], [18]. Ми вважаємо, що застосування активних методів навчання для вирішення проблеми невтручання являє собою ще одну цікаву можливість для дослідження.

**Гнучкість.** Більшість рекомендаційних методів негнучкі в тому сенсі, що вони «жорстко підключені» до систем постачальниками і, отже, підтримують тільки зумовлений і фіксований набір рекомендацій. Тому кінцевий користувач не може налаштувати рекомендації відповідно до своїх потреб в режимі реального часу. Ця проблема була ідентифікована в [19], і для неї була запропонована рекомендаційна мова запитів (Recommendation Query Language) [20]. RQL – це SQL-подібна мова для вираження гнучких рекомендаційних запитів, призначених для користувача. Наприклад, запит «рекомендувати кожному користувачеві з Нью-Йорка кращі три фільми, тривалість яких перевищує дві години», може бути виражена в RQL наступним чином:

```
RECOMMEND Movie TO User
      BASED ON Rating
      SHOW TOP 3
```

```
FROM      MovieRecommender
WHERE     Movie.Length>120
AND      User.City = "NewYork"
```

Крім того, більшість рекомендаційних систем рекомендують тільки окремі елементи для окремих користувачів і не займаються агрегацією. Тим не менш, важливо мати можливість надавати агреговані рекомендації в ряді програм, таких як рекомендації брендів або категорій продуктів для певних сегментів користувачів. Наприклад, рекомендаційна система пов'язана з поїздками може порекомендувати відпочинок у Флориді (категорія продуктів) студентам з північного сходу (призначений для користувача сегмент) під час весняних канікул. Одним із способів підтримки агрегованих рекомендацій є використання підходу до багатовимірних рекомендацій, заснованого на OLAP [21]. Системи на основі OLAP, природно, підтримують ієрархії агрегацій, а початкові підходи до розгортання OLAP-методів в системах рекомендацій представлені в [22]. Проте, потрібно провести роботу для розробки більш повного розуміння того, як використовувати підхід OLAP в рекомендаційних системах, і це уявляє собою цікаву й складну дослідницьку проблему.

**Ефективність рекомендацій.** Проблема розробки хороших показників для оцінки ефективності рекомендацій була широко розглянута в літературі системних рекомендацій. Деякі приклади цієї роботи включають [23], [24]. У більшості літератури про рекомендаційні системи, оцінка ефективності рекомендаційних алгоритмів зазвичай виконується з точки зору метрик охоплення і точності. Охоплення вимірює відсоток елементів, для яких система рекомендацій може робити прогнози [25]. Міри точності можуть бути або статистичними, або застосовуватись для підтримки рішення. Статистичні показники точності в основному порівнюють оцінки рейтингу (наприклад, як визначено в (1)) з фактичними рейтингами  $R$  в матриці Користувач  $\times$  Елемент і включає середню абсолютну помилку (MAE), середньоквадратичну помилку і кореляцію між прогнозами і рейтингами. Міри підтримки прийняття рішень визначають, наскільки добре система рекомендацій може скласти прогнози високорелевантних елементів (наприклад, елементи, які будуть високо оцінені користувачами). Вони включають в себе класичні IR-міри точності (відсоток дійсно «високих» рейтингів серед тих, які були передбачені «високими» в рекомендаційній системі), нагадування/відгук (recall) (відсоток правильно передбачених «високих» значень, рейтинги серед усіх рейтингів, які, відомо є «високими»), показник  $F$  (середнє гармонійне значення точності і відгук) і міра робочої характеристики приймача (Receiver Operating Characteristic, ROC), що демонструє компроміс між істинними позитивними і хибно позитивними показниками в рекомендаційних системах [26].

Незважаючи на популярність, ці емпіричні оцінки мають певні обмеження. Одне з обмежень полягає в тому, що ці заходи зазвичай виконуються на тестових даних, які користувачі вирішили оцінити.

Однак елементи, які користувачі вважають за краще оцінювати, швидше за все, будуть являти собою нерівномірну вибірку, наприклад, користувачі можуть оцінювати в основному ті елементи, які їм подобаються. Іншими словами, результати емпіричної оцінки зазвичай показують тільки те, наскільки точна система на елементах, які користувач вирішив оцінити, тоді як здатність системи правильно оцінювати випадковий елемент (це те, що повинна робити система під час свого нормального життєвого циклу) не перевіряється. Зрозуміло, що це дорого і трудомістко проводити експерименти, які перевіряють якість рекомендацій на неупередженій випадковій вибірці, яка рідко зустрічається. Однак для того, щоб дійсно зрозуміти переваги і обмеження пропонованих рекомендацій, необхідні високоякісні експерименти.

Крім того, хоча важливо є вимірювання точності рекомендацій, технічні заходи, згадані раніше, часто не враховують «корисність» і «якість» рекомендацій. Наприклад, при спостереганні за заявкою в супермаркеті, рекомендуючи очевидні елементи (такі як молоко чи хліб), які споживач купить в будь-якому випадку, буде мати високі показники точності, однак це не буде дуже корисно для споживача. Тому важливо також розробити економічно-орієнтовані міри, такі як показники рентабельності інвестицій (ROI) і цінності життєвого циклу клієнта (LTV) [27]. Розробка і вивчення заходів, спрямованих на усунення обмежень, описаних в цьому розділі, складає цікаву і важливу дослідницьку тему.

**Інші розширення.** Інші важливі проблеми дослідження, які були вивчені в літературі рекомендаційних систем, включають пояснення [12], достовірність [28], масштабованість [4] та конфіденційність [11]. Однак ми не будемо розглядати ці роботи і не будемо обговорювати можливості дослідження в цих областях із-за обмеженості простору.

**SWOT-аналіз отриманих результатів.** Полягає в розподіленні чинників:

- сильних (Strengths)
- і слабких (Weaknesses) сторін проекту,
- можливостей (Opportunities), що відкриваються при його реалізації
- та загроз (Threats), пов'язаних з його здійсненням.

При подальшому розгляді підходів структурована інформація по кожному з напрямків – впливи, слабкі сторони, можливості, загрози – характеризується кількісними значеннями, на основі яких, при застосуванні функцій корисності прораховується потенціал досліджуваного об'єкта по відповідному з напрямків, як підхід, що будується в рамках Conjoint-аналізу.

З появою SWOT-моделі фахівці з обробки аналітичних даних, отримали високоякісний інструмент для інтелектуальної обробки інформації. Загальновідомо, але неструктуровану та несистематизовану інформацію про стан справ у компаніях та про конкурентне оточення, SWOT-аналіз дозволяє сформулювати аналітикам у вигляді логічно погодженої схеми

взаємодії впливів, слабкостей, можливостей і погроз. Сильні сторони компанії сприятимуть забезпеченню просування товарів на ринку та швидшому досягненню стратегічних цілей, у той час як його «слабкості» викликають негативне гальмування. Тут також природно враховувати можливості й загрози зовнішнього середовища, без яких неможливо вірно визначити сценарії розвитку організації. Що в свою чергу сприятиме цінності прогнозів та моніторингу даних, та підвищить конкурентоспроможність гравців ринку.

**Висновки.** В останні десятиліття рекомендаційні системи досягли значного прогресу, коли були запропоновані контентні, колаборативні та гібридні методи, і були розроблені кілька «промислових» систем. Однак, незважаючи на всі ці досягнення, нинішнє покоління рекомендаційних систем, розглянутих в цій статті, як і раніше вимагає подальших поліпшень, щоб зробити рекомендації більш ефективними в більш широкому спектрі застосувань. У цій статті розглянуто різні обмеження існуючих рекомендаційних методів та обговорено можливі розширення, які можуть забезпечити найкращі рекомендації. Ці розширення включають, серед іншого, покращене моделювання користувачів і елементів, включення контекстної інформації в процес рекомендацій, підтримку багатокритеріальних рейтингів і надання більш гнучкого і менш нав'язливого процесу рекомендацій. Ми сподіваємося, що проблеми, представлені в цьому документі, будуть сприяти обговоренню в співтоваристві рекомендаційних систем щодо наступного покоління рекомендаційних технологій.

#### Список літератури

1. Wang Z., Liu X., Li H., Shi J., Rao Y. A Saliency Detection Based Unsupervised Commodity Object Retrieval Scheme. *IEEE Access*. 2018, vol. 6. P. 49902–49912. doi:10.1109/ACCESS.2018.2868139
2. Neves Ana Régia de M., Carvalho Álvaro Marcos G., Ralha Célia G. Agent-based architecture for context-aware and personalized event recommendation. *Expert Systems with Applications*. 2014, vol. 41, issue 2. P. 563–573. <https://doi.org/10.1016/j.eswa.2013.07.081>
3. Wang Shu-Lin, Wu Chun-Yi. Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system. *Expert Systems with Applications*. 2011, vol. 38, issue 9. P. 10831–10838. <https://doi.org/10.1016/j.eswa.2011.02.083>
4. Muhammad Afzal, Syed Imran Ali, Rahman Ali, Maqbool Hussain, Taqdir Ali, Wajahat Ali Khan, Muhammad Bilal Amin, Byeong Ho Kang, Sungyoung Lee. Personalization of wellness recommendations using contextual interpretation. *Expert Systems with Applications*. 2018, vol. 96. P. 506–521. <https://doi.org/10.1016/j.eswa.2017.11.006>
5. Cutting D., Karger D., Pedersen J., Tukey J. A Cluster-based Approach to Browsing Large Document Collections. *Semantic scholar*. URL: <https://pdfs.semanticscholar.org/1134/3448f8a817fa391e3a7897a95f975ad2873a.pdf> (дата звернення: 10.10.2018).
6. Yaohao Peng, Pedro Henrique Melo Albuquerque, Jader Martins Camboim de Sá, Ana Julia Akaishi Padula, Mariana Rosa Montenegro. The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression. *Expert Systems with Applications*. 2018, vol. 97. P. 177–192. <https://doi.org/10.1016/j.eswa.2017.12.004>
7. Sasvári Tamás. *Language Detector class (C#)*. URL: [http://idsyst.hu/development/language\\_detector.html](http://idsyst.hu/development/language_detector.html) (дата звернення: 01.10.2018).



8. Бутова Р. К., Назарова Г. Н., Павленко Л. А. та ін. *Інформаційні системи і технології в економіці: посіб. для студ. вищих навч. закл.* Київ: Академія, 2002. 544 с
9. Апопій В. В., Олексин І. І., Шутовська Н. О., Футало Т. В. *Організація і технологія надання послуг: навч. посіб.* Київ: Академія, 2016. 312 с.
10. Федуллова Л. І. *Інноваційна економіка: підручник.* Київ: Либідь, 2006. 480 с.
11. Гець В. Характер перехідних процесів до економіки знань. *Економіка України.* 2004. № 5. С. 4–13.
12. Nilashi M., Esfahani M. D., Roudbaraki M. Z., Ramayah, T., Ibrahim O. A multi-criteria collaborative filtering recommender system using clustering and regression techniques. *Journal of Soft Computing and Decision Support Systems.* 2016, vol. 3, no. 5. P. 24–30.
13. *Економічна правда.* 2018, 8 грудня. URL: <https://www.epravda.com.ua/> (дата звернення: 10.10.2018).
14. Ricci F., Rokach L., Shapira B. Recommender systems: introduction and challenges. *Recommender systems handbook.* Boston: Springer, 2015. P. 1–34.
15. Sivaraman S., Trivedi M. M. Active learning for on-road vehicle detection: A comparative study. *Machine vision and applications.* 2014. Vol. 25, no. 3. P. 599–611.
16. Rubens N., Elahi M., Sugiyama M., Kaplan D. Active learning in recommender systems. *Recommender systems handbook.* 2015, Boston: Springer, 2015. P. 809–846.
17. Buczak A. L., Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials.* 2016, vol. 18, no. 2. P. 1153–1176.
18. Ulissi Z. W., Medford A. J., Bligaard T., Nørskov J. K. To address surface reaction network complexity using scaling relations machine learning and DFT calculations. *Nature communications.* 2017, vol. 8. P. 14621–14627.
19. Hwangbo H., Kim Y. S., Cha K. J. Recommendation system development for fashion retail e-commerce. *Electronic Commerce Research and Applications.* 2018, vol. 28. P. 94–101.
20. Wongsuphasawat K., Moritz D., Anand A., Mackinlay J., Howe B., Heer J. Towards a general-purpose query language for visualization recommendation. *Proceedings of the Workshop on Human-In-the-Loop Data Analytics.* URL: <https://www.domoritz.de/papers/2016-Compasql-HILDA.pdf> (дата звернення: 01.10.2018).
21. Yashchuk D. Y., Golub, B. L. Research on the Use of OLAP Technologies in Management Tasks. *International Conference on Theory and Applications of Fuzzy Systems and Soft Computing.* Cham: Springer, 2018. P. 683–691.
22. Gallinucci, E., Golfarelli, M., Rizzi, S. *Variety-Aware OLAP of Document-Oriented Databases.* 2018.
23. Kluver D., Ekstrand M. D., Konstan J. A. (). Rating-based collaborative filtering: algorithms and evaluation. *Social Information Access.* Springer, Cham, 2018. P. 344–390.
24. Wang D., Liang Y., Xu D., Feng X., Guan R. A content-based recommender system for computer science publications. *Knowledge-Based Systems.* 2018, vol. 157. P. 1–9.
25. Аулін В. В., Голуб Д. В. Методи оцінки і аналізу надійності автомобільних транспортних систем. *Збірник тез доповідей VI Міжнародної науково-технічної конференції молодих учених та студентів «Актуальні задачі сучасних технологій».* 2017. Т. 3. С. 14–15.
26. Leris D., Sein-Echaluce M. L., Hernández, M., Bueno C. Validation of indicators for implementing an adaptive platform for MOOCs. *Computers in Human Behavior.* 2017, vol. 72. P. 783–795.
27. Chaikovska M., Chaikovskiy M. Strategies for implementation of affiliate-projects in marketing activity. *Scientific Journal of Polonia University.* 2018, vol. 27, no. 2. P. 18–25.
28. Wang S., Zheng Z., Wu Z., Lyu M. R., Yang F. Reputation measurement and malicious feedback rating prevention in web service recommendation systems. *IEEE Transactions on Services Computing.* 2015, vol. 8, no. 5. P. 755–767.
29. Neves Ana Régia de M., Carvalho Álvaro Marcos G., Ralha Célia G. Agent-based architecture for context-aware and personalized event recommendation. *Expert Systems with Applications.* 2014, vol. 41, issue 2, pp. 563–573. <https://doi.org/10.1016/j.eswa.2013.07.081>
30. Wang Shu-Lin, Wu Chun-Yi. Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system. *Expert Systems with Applications,* 2011, vol. 38, issue 9, pp. 10831–10838. <https://doi.org/10.1016/j.eswa.2011.02.083>
31. Muhammad Afzal, Syed Imran Ali, Rahman Ali, Maqbool Hussain, Taqdir Ali, Wajahat Ali Khan, Muhammad Bilal Amin, Byeong Ho Kang, Sungyoung Lee. Personalization of wellness recommendations using contextual interpretation. *Expert Systems with Applications.* 2018, vol. 96, pp. 506–521. <https://doi.org/10.1016/j.eswa.2017.11.006>
32. Cutting D., Karger D., Pedersen J., Tukey J. *A Cluster-based Approach to Browsing Large Document Collections. Semantic scholar.* Available at: <https://pdfs.semanticscholar.org/1134/3448f8a817fa391e3a7897a95f975ad2873a.pdf> (accessed 10.10.2018).
33. Yaohao Peng, Pedro Henrique Melo Albuquerque, Jader Martins Camboim de Sá, Ana Julia Akaishi Padula, Mariana Rosa Montenegro. The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression. *Expert Systems with Applications.* 2018, vol. 97, pp. 177–192. <https://doi.org/10.1016/j.eswa.2017.12.004>
34. Sasvári Tamás. *Language Detector class (C#).* Available at: [http://idsyst.hu/development/language\\_detector.html](http://idsyst.hu/development/language_detector.html) (accessed 01.10.2018).
35. Butova R. K., Nazarov G. N., Pavlenko L. A. and others. *Інформаційні системи і технології в економіці: посіб. для студ. вищих навч. закл.* [Information systems and technologies in the economy: manual. for studio higher educators shut up]. Kiev, Academy Publ., 2002. 544 p.
36. Apopia V. V., Olexin I. I., Shutovskaya N. O., Futalo T. V. *Організація і технологія надання послуг: навчальний посібник* [Organization and technology of providing services: Teaching manual]. Kiev, Academy Publ., 2016. 312 pp.
37. Fedulova L. I. *Інноваційна економіка: підручник* [Innovative economy: textbook]. Kiev, Lybid Publ., 2006. 480 p.
38. Geets V. Charakter perehidnyh procesiv do ekonomiky znan' [Character of Transitional Processes to the Knowledge Economy]. *Економіка України* [Economics of Ukraine]. 2004, no. 5, pp. 4–13.
39. Nilashi M., Esfahani M. D., Roudbaraki M. Z., Ramayah, T., Ibrahim O. A multi-criteria collaborative filtering recommender system using clustering and regression techniques. *Journal of Soft Computing and Decision Support Systems.* 2016, vol. 3, no. 5, pp. 24–30.
40. *Економічна правда* [Economic truth]. 2018, December 18. [Electronic resource]. Available at: <https://www.epravda.com.ua/> (accessed 01.10.2018)
41. Ricci F., Rokach L., Shapira B. Recommender systems: introduction and challenges. *Recommender systems handbook.* Boston, Springer, 2015, pp. 1–34.
42. Sivaraman S., Trivedi M. M. Active learning for on-road vehicle detection: A comparative study. *Machine vision and applications.* 2014, vol. 25, no. 3, pp. 599–611.
43. Rubens N., Elahi M., Sugiyama M., Kaplan D. Active learning in recommender systems. *Recommender systems handbook.* Boston, Springer, 2015, pp. 809–846.
44. Buczak A. L., Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials.* 2016, vol. 18, no. 2, pp. 1153–1176.
45. Ulissi Z. W., Medford A. J., Bligaard T., Nørskov J. K. To address surface reaction network complexity using scaling relations machine learning and DFT calculations. *Nature communications.* 2017, vol. 8, pp. 14621–14627.
46. Hwangbo H., Kim Y. S., Cha K. J. Recommendation system development for fashion retail e-commerce. *Electronic Commerce Research and Applications.* 2018, vol. 28, pp. 94–101.
47. Wongsuphasawat K., Moritz D., Anand A., Mackinlay J., Howe B., Heer J. Towards a general-purpose query language for visualization recommendation. *Proceedings of the Workshop on Human-In-the-*

## References (transliterated)

1. Wang Z., Liu X., Li H., Shi J., Rao Y. A Salient Detection Based Unsupervised Commodity Object Retrieval Scheme. *IEEE Access.* 2018, vol. 6, pp. 49902–49912. doi: 10.1109/ACCESS.2018.2868139.

- Loop Data Analytics. Available at: <https://www.domoritz.de/papers/2016-Compassql-HILDA.pdf> (accessed 01.10.2018).
21. Yashchuk D. Y., Golub, B. L. Research on the Use of OLAP Technologies in Management Tasks. *International Conference on Theory and Applications of Fuzzy Systems and Soft Computing*. Cham, Springer, 2018, pp. 683–691.
  22. Gallinucci, E., Golfarelli, M., Rizzi, S. *Variety-Aware OLAP of Document-Oriented Databases*. 2018.
  23. Kluver D., Ekstrand M. D., Konstan J. A. (). Rating-based collaborative filtering: algorithms and evaluation. *Social Information Access*. Springer, Cham, 2018, pp. 344–390.
  24. Wang D., Liang Y., Xu D., Feng X., Guan R. A content-based recommender system for computer science publications. *Knowledge-Based Systems*. 2018, vol. 157, pp. 1–9.
  25. Aulin V. V., Holub D. V. Metody otsinky i analizu nadiynosti avtomobil'nykh transportnykh system [Methods of estimation and analysis of reliability of automobile transport systems]. *Zbirnyk tez dopovidej VI Mizhnarodnoi' naukovykh-tehnichnoi' konferencii' molodykh uchenykh ta studentiv „Aktual'ni zadachi suchasnykh tehnologij“* [Proc. of VI International Science and Technology and Conference of Young Researchers and Students „Current problems of modern technologies“]. 2017, vol. 3, pp. 14–15.
  26. Leris D., Sein-Echaluce M. L., Hernández, M., Bueno C. Validation of indicators for implementing an adaptive platform for MOOCs. *Computers in Human Behavior*. 2017, vol. 72, pp. 783–795.
  27. Chaikovska M., Chaikovskiy M. Strategies for implementation of affiliate-projects in marketing activity. *Scientific Journal of Polonia University*. 2018, vol. 27, no. 2, pp. 18–25.
  28. Wang S., Zheng Z., Wu Z., Lyu M. R., Yang F. Reputation measurement and malicious feedback rating prevention in web service recommendation systems. *IEEE Transactions on Services Computing*. 2015, vol. 8, no. 5, pp. 755–767.

Надійшла (received) 19.11.2018

## Відомості про авторів / Сведения об авторах / About the Authors

**Чередніченко Ольга Юрївна (Чередниченко Ольга Юрьевна, Cherednichenko Olga Yuryevna)** – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри Програмної інженерії та інформаційних технологій управління; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-9391-5220>; e-mail: [olha.cherednichenko@gmail.com](mailto:olha.cherednichenko@gmail.com)

**Іващенко Оксана Віталіївна (Иващенко Оксана Витальевна, Ivashchenko Oksana Vitaliivna)** – Національний технічний університет «Харківський політехнічний інститут», аспірант; м. Харків, Україна, ORCID: <https://orcid.org/0000-0003-3636-3914>; e-mail: [Oksana\\_ivashchenko@ukr.net](mailto:Oksana_ivashchenko@ukr.net)

**Гонтар Юлія Миколаївна (Гонтарь Юлия Николаевна, Gontar Yulia Mykolaivna)** – Національний технічний університет «Харківський політехнічний інститут», аспірант; м. Харків, Україна, ORCID: <https://orcid.org/0000-0001-2345-6789>; e-mail: [gontaryn@gmail.com](mailto:gontaryn@gmail.com)

**Ворона Борис Михайлович (Ворона Борис Михайлович, Vorona Borys Mykhailovych)** – Національний технічний університет «Харківський політехнічний інститут», студент; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-9649-0264>; e-mail: [borysvorona@gmail.com](mailto:borysvorona@gmail.com)

УДК 004.891.3

DOI: 10.20998/2079-0023.2018.44.11

**В. О. ЛЕЩИНСЬКИЙ, І. О. ЛЕЩИНСЬКА****ЗАСТОСУВАННЯ АКТИВНОГО НАВЧАННЯ В СИТУАЦІЇ ЦИКЛІЧНОГО ХОЛОДНОГО СТАРТУ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ**

Досліджено проблему побудови рекомендацій для систем електронної комерції в умовах циклічного холодного старту. Дана проблема виникає при постійній зміні інтересів користувачів протягом строку використання рекомендаційної підсистеми. Існуючі підходи до формування рекомендацій в умовах холодного старту засновані на поступовому накопиченні інформації про споживача і тому не забезпечують релевантних рекомендацій у випадку циклічних змін у їх діяльності та інтересах. Для вирішення цієї проблеми пропонується враховувати аспект зміни інтересів користувачів щодо товарів та послуг в часі. Мета даної роботи полягає в уточненні принципів активного навчання для побудови рекомендацій в умовах зміни інтересів споживачів з тим, щоб забезпечити поетапне уточнення персональних рекомендацій «холодним» споживачам. Отримані результати містять у собі деталізовану задачу формування рекомендацій, та уточнені принципи активного навчання. Виділено ключову особливість циклічного холодного старту в рекомендаційній системі, що полягає в обмеженості періоду, протягом якого може бути доповнена та уточнена інформація про клієнта системи електронної комерції. Зазначена особливість обумовлює потребу в врахуванні аспекту часу при формуванні рекомендацій щодо вибору товарів та послуг. Сформульовано задачу формування рекомендацій в умовах циклічного холодного старту як задачу ітеративного доповнення та уточнення даних нового «холодного» користувача патернами найбільш розповсюджених циклів поведінки споживачів з подальшим використанням колаборативної фільтрації уточнених даних для формування рекомендацій. Доповнено принципи застосування активного навчання для умов циклічного холодного старту на основі використання типових послідовностей дій користувача у часі. Вказані принципи дозволяють коригувати вхідні дані для «холодного» користувача за допомогою евристичної стратегії, що враховує зміни у патернах поведінки споживачів. Патерни поведінки відображають циклічні зміни інтересу споживачів щодо товарів та послуг, що пропонує система електронної комерції.

**Ключові слова:** рекомендаційні системи, ранжування результатів; холодний старт, колаборативна фільтрація, формування рекомендацій.

**В. А. ЛЕЩИНСКИЙ, И. А. ЛЕЩИНСКАЯ****ПРИМЕНЕНИЕ АКТИВНОГО ОБУЧЕНИЯ В СИТУАЦИИ ЦИКЛИЧЕСКОГО ХОЛОДНОГО СТАРТА РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ**

Исследована проблема построения рекомендаций для систем электронной коммерции в условиях циклического холодного старта. Данная

© В. О. Лещинський, І. О. Лещинська, 2018