

УДК 004.891

## Розробка гібридної рекомендаційної системи

А.О. Белан, Л.В. Васильєва

**Белан****Анжеліка Олександрівна***студентка**Харківський Національний Університет імені В. Н. Каразіна,**майдан Свободи 6, Харків, 61022**e-mail: [anzhelika.bielan@student.karazin.ua](mailto:anzhelika.bielan@student.karazin.ua)**<https://orcid.org/0009-0006-5473-0430>***Васильєва****Лариса Валентинівна***к.б.н., доцент;**Харківський Національний Університет імені В. Н. Каразіна,**майдан Свободи 6, Харків, 61022**e-mail: [lvvasilieva@karazin.ua](mailto:lvvasilieva@karazin.ua)**<https://orcid.org/0000-0001-7926-3062>*

Сучасний інтернет і веб-сервіси переповнені великою кількістю інформації, що стає складнішим для користувачів. Системи рекомендацій мають на меті вирішити проблему перевантаження інформацією, одночасно персоналізують досвід користувача, надають точні персоналізовані рекомендації відповідно до їхніх уподобань. Головною метою цієї роботи є реалізація системи рекомендацій для складної предметної сфери та алгоритму її впровадження для предметної області пошуку закладів послуг. У статті коротко описані основні підходи та алгоритми, які використовуються в системах рекомендацій, а також висвітлені області застосування, переваги та недоліки кожного з них. У роботі наведено опис реалізації системи рекомендацій для складної предметної сфери та алгоритму її впровадження для предметної області пошуку закладів сфери послуг. Особливістю системи, що розробляється, є важливість даних про геолокацію для формування рекомендацій. Розроблено алгоритм гібридної системи рекомендацій, що об'єднує в собі підходи до розробки систем рекомендацій на основі знань та фільтрації на основі вмісту. Перевагами описаного алгоритму є відсутність необхідності зберігати та використовувати у розрахунках інформацію про попередні сеанси користувачів, вирішення проблеми холодного запуску, формування актуальних рекомендацій для поточного сеансу користувача та вирішення проблеми визначення місцезнаходження для рекомендованого закладу. Запропонований підхід може бути використано для розробки та впровадження рекомендаційного алгоритму у складних предметних сферах, де дані про геолокацію важливі для надання рекомендацій.

**Ключові слова:** система рекомендацій, рекомендаційний алгоритм, гібридна система рекомендацій, фільтрація за вмістом, рекомендації на основі знань.

**Як цитувати:** Белан А.О., Васильєва Л.В. Розробка гібридної рекомендаційної системи. *Вісник Харківського національного університету імені В. Н. Каразіна, сер. «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління»*. 2023. вип. 57. С.22-31.

<https://doi.org/10.26565/2304-6201-2023-57-02>

### 1 Вступ

У сучасному світі з динамічним розвитком другого покоління мережевих сервісів інтернету WEB 2.0 кожен має доступ до надлишково великої кількості інформації. Користувачам все складніше впоратися з цим потоком даних та виділити для себе важливу інформацію. На інтернет-платформах електронної комерції, де пропонуються мільйони товарів, звичайному споживачеві складно охопити всі можливості, і це може призвести до перенасичення інформацією. Системи рекомендацій мають на меті вирішити цю проблему перевантаження інформацією, одночасно персоналізують досвід користувача, надають точні персоналізовані рекомендації щодо предметів/продуктів користувачам відповідно до їхніх уподобань. Метою системи рекомендацій є передбачення, чи буде продукт корисним для користувача на основі відомої про нього інформації.

Застосування систем рекомендацій (СР) в останні роки постійно зростає і вони використовуються у сферах роздрібною торгівлі та електронної комерції, таких як eBay, Amazon та Netflix. Ці компанії отримують величезну кількість даних користувачів і використовують СР для задоволення потреб користувачів і покращення бізнес-процесів. За оцінками McKinsey, 35% виручки Amazon і 75% Netflix припадає саме на рекомендовані товари, і цей відсоток, швидше за все, зростатиме [1].

На сьогоднішній день сфера розробки CR переживає значний розвиток. Такі системи використовуються поряд та разом з генеративним штучним інтелектом, що стрімко розвивається саме сьогодні. Однією з ключових тенденцій є використання машинного навчання і глибокого навчання, щоб створювати більш точні та ефективні рекомендації для користувачів [2]. Використання засобів обробки природної мови, а також збільшення вимог до якості даних і етичності алгоритмів є також надзвичайно важливими аспектами в розробці CR.

Персоналізація також стала важливим аспектом розробки CR, де моделі адаптуються до індивідуальних потреб і смаків користувачів на основі їхнього минулого взаємодії. CR використовують різноманітні дані користувачів для створення більш точних і персоналізованих рекомендацій. Ці дані включають в себе історію взаємодії користувачів на платформі, такі як перегляди, оцінки, покупки, які допомагають розуміти їхні вподобання та інтереси [3]. Демографічні дані, такі як вік, стать і місце проживання, допомагають категоризувати користувачів. Дані про геолокацію важливі для рекомендацій, пов'язаних з місцевими послугами [4]. Розробники продовжують експериментувати з новими методами і техніками, а також враховувати зростаючий інтерес до прозорості та етичних питань в роботі з даними.

## **2 Категорії систем рекомендацій**

Системи рекомендацій можна класифікувати за кількома основними категоріями, залежно від способу їхньої роботи та особливостей алгоритмів, що використовуються [5].

### **2.1. Системи рекомендацій на основі вмісту (Content-Based Recommendation Systems)**

Підходи, спрямовані на створення CR на основі вмісту, мають за мету побудувати профіль користувача, щоб передбачити рейтинги об'єктів, які користувач ще не оцінив. Методи, базовані на вмісті, використовують для цього теги та ключові слова. Оцінка корисності фільтрації на основі вмісту зазвичай проводиться за допомогою евристичних метрик, таких як косинус подібності. Цей підхід до рекомендацій може бути застосованим в ситуаціях, коли характеристики об'єктів можна легко визначити. Однак, він неефективний у випадках, коли вимагається складна і вручну введена інформація про об'єкти, особливо в разі великої кількості нових продуктів щодня.

Фільтрація на основі вмісту не потребує даних інших користувачів, оскільки рекомендації індивідуальні для кожного користувача. Це робить цей метод масштабованим для великої кількості користувачів і не залежним від їхніх дій. Системи на основі вмісту вимагають великих обсягів даних та глибоких знань у предметній області для точних рекомендацій. Крім того, вони мають обмежену здатність до виявлення нових об'єктів або інноваційних рекомендацій, оскільки вони опираються на характеристики інших оцінених об'єктів та профілі користувачів [7].

### **2.2. Системи рекомендацій на основі колаборативної фільтрації (Collaborative Filtering Recommendation Systems)**

Колаборативна фільтрація це метод оцінки продуктів на основі рейтингів, які надають користувачі у історичних даних. Цей метод створює базу даних уподобань користувачів до об'єктів, шляхом виявлення сусідніх активних користувачів за подібними перевагами в покупках.

Методи колаборативної фільтрації поділяються на дві основні категорії: фільтрацію на основі схожості елементів (продуктів) та фільтрацію на основі схожості користувачів [6]. Методи, орієнтовані на користувача, складаються з двох основних кроків для прогнозування рейтингів товарів для конкретного користувача: пошук користувачів, подібних до цільового користувача та отримання оцінок від користувачів, схожих на активного користувача. Ці оцінки використовуються для генерації рекомендацій.

Існує багато алгоритмів колаборативної фільтрації, які обчислюють схожість між користувачами. Зазвичай використовуються міри подібності, такі як середньоквадратична різниця, кореляція Пірсона, косинус подібності, кореляція Спірмена та скоригована косинус подібність [6]. Колаборативна фільтрація є популярним вибором для CR і не вимагає глибокого розуміння домену, оскільки вбудовується та розраховується автоматично.

CR на основі колаборативної фільтрації мають свої недоліки, включаючи проблеми холодного запуску через нестачу даних для нових користувачів і об'єктів, розрідженість даних оцінок у великих системах, складності масштабування при зростанні кількості користувачів та об'єктів.

### **2.3. Системи рекомендацій на основі демографічних даних (Demographic-Based Recommendation Systems)**

Результати різних кількісних досліджень демонструють, що методи спільної фільтрації можуть бути покращені за допомогою демографічної кореляції. SR на основі демографії здатні генерувати рекомендації, класифікуючи користувачів за їхніми демографічними характеристиками. Ці системи використовують атрибути користувачів, такі як вік, стать, мова та інші, для надання рекомендацій [8].

Основною перевагою демографічних фільтрів SR є швидкість та простота отримання результатів на основі обмеженого обсягу спостережень. SR на основі демографії спрямовані на вирішення проблем масштабованості та холодного запуску. Вони стають особливо корисними тоді, коли обмежена кількість інформації про продукт. Проте ці методи мають свої недоліки, такі як проблеми зі збором та збереженням конфіденційної інформації користувачів та недостатня гнучкість у врахуванні зміни уподобань користувачів, що може вплинути на стабільність та ефективність рекомендаційної системи.

### **2.4. Системи рекомендацій на основі корисних програм (Utility-Based Recommendation Systems)**

SR на основі корисних програм пропонують рекомендації на підставі розробки корисної моделі для кожного елемента, спрямованої на задоволення потреб користувача. Ці системи створюють багатоатрибутні функції корисності для користувачів та рекомендують елементи на основі явно обчисленої корисності кожного з них. SR на основі утиліт важливі, оскільки вони можуть враховувати не лише характеристики продуктів, але й інші атрибути, такі як доступність товарів та надійність постачальників. Вони здатні генерувати обчислення корисності в реальному часі, що дозволяє оцінити стан товару для користувача. Системи на основі утиліт не базуються на довгострокових узагальненнях про користувачів, а враховують актуальні потреби користувачів та доступні варіанти. Недоліком таких систем є те, що система може сховати рекомендації від користувачів, навіть якщо вони відповідають вподобанням цього користувача, якщо описові характеристики товарів обмежені [9].

### **2.5. Системи рекомендацій на основі знань (Knowledge-Based Recommendation Systems)**

SR на основі знань використовують явні знання про продукти та користувачів для створення критерію на основі знань з метою розробки рекомендацій. На відміну від інших SR, вони не залежать від великої кількості статистичних даних про конкретні оцінені товари чи конкретних користувачів. Системи на основі знань мають декілька переваг. Наприклад, вони можуть уникнути проблеми нарощування, що часто виникає в методах машинного навчання для рекомендацій. Зазвичай зразкові системи не можуть навчатися, доки користувач не надає багато оцінок. SR на основі знань уникають цієї проблеми, оскільки їх рекомендації не залежать від історії оцінок користувача. Крім того, вони не вимагають збирання конкретних даних про користувача, оскільки рекомендації також не підпорядковані індивідуальним смакам користувача. З цих причин системи на основі знань цінні як самостійні SR і можуть доповнювати інші типи.

Одним з основних недоліків систем на основі знань є потенційна обмеженість в питанні отримання знань, яка виникає внаслідок явного визначення рекомендаційних знань. Процес отримання знань передбачає створення правил і вимог, які необхідні для системи на основі знань, і це здійснюється шляхом надання знань на основі правил, об'єктів і онтологій, побудованих на фреймах. Теорії Бейтсона були використані для керування процесом подальшого вивчення отримання знань [10].

### **2.6. Гібридні системи рекомендацій (Hybrid-Based Recommendation Systems)**

Гібридні SR представляють собою підхід до розробки систем, який поєднує в собі різні методи та підходи з метою покращити якість та ефективність рекомендацій. Ці системи поєднують в собі два або більше основних типи SR, такі як колаборативна фільтрація, фільтрація на основі вмісту, системи на основі корисних програм, базовані на знаннях та інші. Головна мета гібридних систем це компенсація обмежень та недоліків кожного окремого методу, створюючи стійку та більш точну SR [11].

Однією з ключових переваг гібридних систем є здатність знаходити підхід до різних видів даних та користувачів. Наприклад, якщо один метод заснований на спільній фільтрації добре

працює для користувачів з великою історією оцінок, а інший метод на основі вмісту вдається для нових користувачів, гібридна система може комбінувати їхні сильні сторони для надання кращих рекомендацій незалежно від користувача.

Створення гібридних систем може бути складним завданням, оскільки потребує інтеграції різних методів та визначення оптимального способу комбінування їх результатів. Важливо також вирішити питання вагомості різних компонентів гібридної системи та розробити ефективні алгоритми комбінування рекомендацій. Гібридні CR стають є найбільш популярними завдяки їхній здатності покращувати точність та роботу CR, особливо в умовах великого обсягу даних та різноманітності користувачів.

### 3 Практична реалізація системи рекомендацій

Головною метою цієї роботи є реалізація системи рекомендацій для складної предметної сфери та алгоритму її впровадження для предметної області пошуку закладів послуг. Особливістю системи, що розроблюється, є можливість отримання даних про місцезнаходження користувача та доступність геоданих закладів послуг, які зберігаються в базі даних системи. Для дослідження було запропоновано виконати аналіз можливих шляхів впровадження рекомендацій у систему, виділити основні аспекти та ідеї на основі яких буде розроблена CR.

Важливим аспектом розробки алгоритму рекомендацій є визначення етапу взаємодії з системою, коли користувачу будуть надаватись рекомендації. Послідовність дій, які користувач може виконати в системі зображено на діаграмі діяльності (рис. 1). Для даної предметної області впровадження рекомендацій може виконуватись на етапі початкового перегляду доступних закладів, під час вибору параметрів для пошуку та при перегляді деталей про обраний заклад. Причому, підходи до розробки алгоритму рекомендацій можуть сильно різнитися у залежності від етапу взаємодії користувача з системою, на якому будуть надаватись рекомендації.

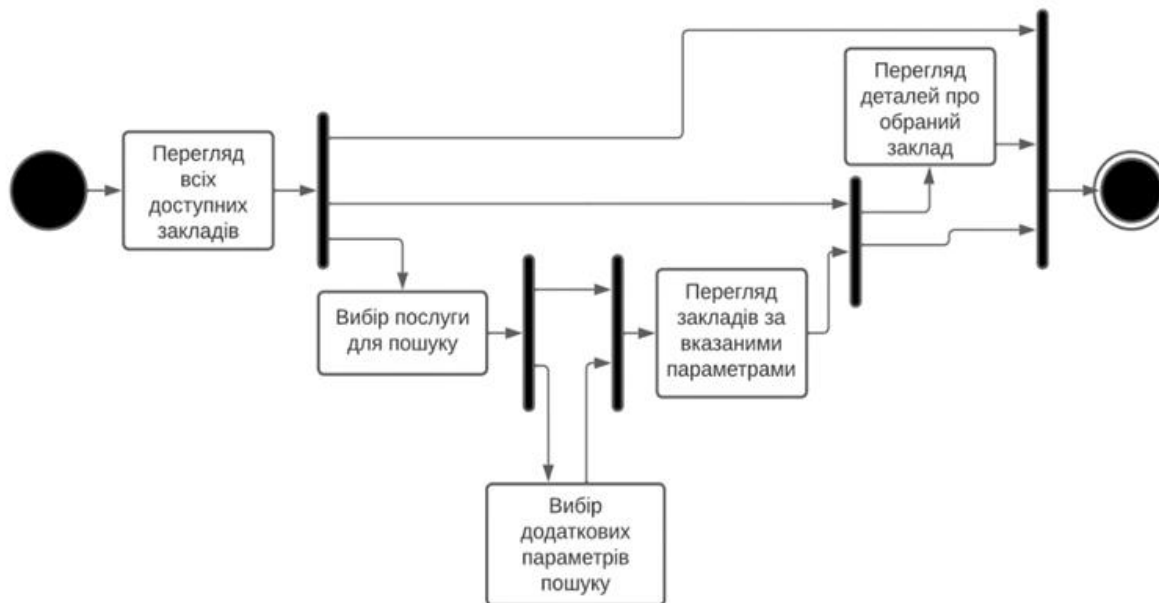


Рис.1 Діаграма активності

На етапі входу в систему та можливості перегляду всіх доступних закладів на карті користувачу можуть пропонуватися заклади за його місцезнаходженням, які надають послуги, що були оцінені користувачем у попередніх сесіях. У випадку, коли користувач вперше взаємодіє з системою, йому можуть пропонуватися найбільш популярні заклади та послуги, які вони надають, за його місцезнаходженням.

Після вибору користувачем послуги та додаткових параметри для пошуку найбільш доцільно рекомендувати заклади та послуги, що найкраще відповідають заданим параметрам, у тому числі тоді, коли в базі даних системи відсутні заклади та послуги, які повністю відповідають вказаними параметрам.

При наданні рекомендацій на етапі перегляду користувачем деталей про заклад, що надає шукану ним послугу, мають враховуватися параметри закладу та послуги. У цьому випадку для формування рекомендацій, на відміну від попередніх випадків, більше значення мають дані про місцезнаходження переглянутого закладу, ніж про місцезнаходження користувача.

Для дослідження пропонується розглянути розробку алгоритму роботи СР, впроваджену на етапі пошуку закладів послуг за вказаними параметрами та перегляду деталей про обрані заклади. Розроблений алгоритм рекомендацій має вирішувати проблеми холодного запуску та відсутності в базі даних системи об'єктів, що повністю відповідають вказаним параметрам пошуку, а також проблему визначення місцезнаходження для рекомендації.

З проаналізованих основних підходів до розробки рекомендаційних систем можна визначити, що вдало побудовані гібридні СР справляють з задачею надання рекомендацій найкраще. Отже, для вирішення проблеми відсутності в базі даних системи об'єктів, що повністю відповідають вказаним параметрам пошуку, застосовується підхід до розробки СР на основі вмісту, який враховує та аналізує характеристики закладу та послуги, які користувач шукає або переглядає. На першому етапі роботи алгоритму користувачу пропонуються об'єкти, які є найбільш схожими до так званої «ідеальної точки». Ідеальною точкою буде вважатися об'єкт із характеристиками сформованими на основі параметрів пошуку та місцезнаходження користувача або ж об'єкт, який користувач переглядає в даний момент. Для визначення «схожості» між об'єктами використовуються показники подібності для порівняння атрибутів цих об'єктів. Визначаються дві основні речі для кожного атрибута: його вага у відношенні до інших атрибутів і формула або метод, за допомогою якого оцінюється схожість між двома значеннями цього атрибута.

Для вирішення проблеми холодного запуску застосовується підхід розробки СР на основі знань, який полягає в формуванні списку вподобань користувача в рамках поточного сеансу застосування параметрів фільтрації. Так звана ідея «діалогу» застосовується на другому етапі роботи алгоритму та полягає в тому, що після визначення параметрів пошуку та перегляду певної кількості пропозицій рекомендації будуть базуватися на групі з деякої кількості останніх переглянутих закладів та послуг, які вони надають.

Крім того, для рекомендаційного алгоритму вводяться обмеження, які зменшують простір об'єктів, що можуть бути рекомендовані користувачу. Оскільки основною метою системи є пошук користувачем закладів, які надають послуги, що його цікавлять, то не доцільно рекомендувати йому заклади, які не надають послуг цієї сфери. Відповідний підхід стосується і міста, де користувач шукає заклади, що надають відповідну послугу.

Математичні основи даного алгоритму включають у себе наступне. Для кількісної оцінки схожості деякого об'єкта з еталонним використовується метод зваженої суми критеріїв. Кожен об'єкт представлено вектором критеріїв:

$$K = \{k_1, k_2, k_3, \dots, k_n\} \quad (1)$$

де  $k_i$  - кількісна оцінка схожості  $i$ -ої характеристики об'єкта з  $i$ -ою характеристикою еталонного елемента.

Відносна важливість кожної з характеристик виражена вектором вагових коефіцієнтів, сума яких рівна одиниці:

$$W = \{w_1, w_2, w_3, \dots, w_n\} \quad (2)$$

де  $w_i$  - відносна важливість  $i$ -ої характеристики об'єкта.

Результуюче значення схожості об'єкту розраховується за формулою:

$$Score = \sum w_i \cdot k_i \quad (3)$$

Для оцінки на основі групи елементів, оцінки (3) нового елемента, що розраховуються відносно кожного елемента у межах групи переглянутих, підсумовуються між собою для отримання кінцевого результату:

$$Score^* = \sum_{i=1}^N Score_i \quad (4)$$

де  $N$  – кількість переглянутих елементів.

Для оцінки схожості характеристик об'єкта з характеристиками еталонного об'єкта, представленими у вигляді інтервальних оцінок, використовуються функції затухання (або функції розподілу). Ці функції оцінюють певну величину, зменшуючи своє значення залежно від збільшення відстані між значенням цієї величини та певним заданим джерелом (центром мас). Ця оцінка подібна до запиту щодо включення деякого значення в заданий діапазон, з подальшим призначенням кількісної оцінки цьому значенню в залежності від близькості до центра діапазону.

Існують три основні форми функцій затухання (розподілу), які використовуються:

- Гауссіан (за функцією Гауса);
- експоненціальний розподіл;
- лінійний розподіл.

У розробці рекомендаційної системи використовується гаусівська залежність, оскільки лінійна залежність є занадто простою та приводить до однакового зменшення оцінок, при віддаленні від центрального значення, а експоненціальна залежність має дуже різкий графік зменшення значень біля центру та менш різкий за його межами. У даному випадку потрібна зворотна залежність, яку надає розподіл Гауса.

Затухання для характеристики об'єкта визначається наступним чином:

$$K_i = \exp \left( - \frac{(|value_i - origin| - offset)^2}{2 \cdot \left( - \frac{scale^2}{2 \cdot \ln(decay)} \right)} \right) \quad (4)$$

де  $K_i$  – кількісна оцінка схожості  $i$ -тої характеристики об'єкта з  $i$ -тою характеристикою еталонного об'єкта, приймає значення в діапазоні  $[0;1]$ ;

$value_i$  – значення  $i$ -тої характеристики об'єкта;

$origin$  – значення  $i$ -тої характеристики еталонного об'єкта;

$offset$  – відстань, на якій значення характеристики буде вважатися рівним еталонній;

$scale$  – відстань від центру (+ зміщення) яка визначає діапазон прийнятної (допустимої) схожості;

$decay$  – значення схожості в яке оцінюється значення характеристики на відстані  $scale$  від  $origin$ .

Для характеристик, які представлені на номінальних шкалах, ми можемо визначити лише подібність або різницю між об'єктами. Таким чином, розрахунок подібності для таких характеристик буде виглядати наступним чином:

–  $K_i = 1$ , якщо значення відповідних характеристик однакові, що вказує на повне співпадіння значень;

–  $K_i = 0$ , якщо значення відповідних характеристик різні, що вказує на повну відсутність співпадіння, і це оцінюється нульовим балом.

Якщо значення ж певної характеристики відсутнє, то оцінка схожості цієї характеристики між об'єктом і еталонним об'єктом буде  $K_i = 0$ .

Блок-схему описаного алгоритму зображено на рисунку 2. Відображення результатів роботи СР відбувається після застосування користувачем фільтрації, що передбачає вибір послуги для пошуку та інших параметрів, а також при перегляді детальної інформації про заклади, які є результатом пошуку.

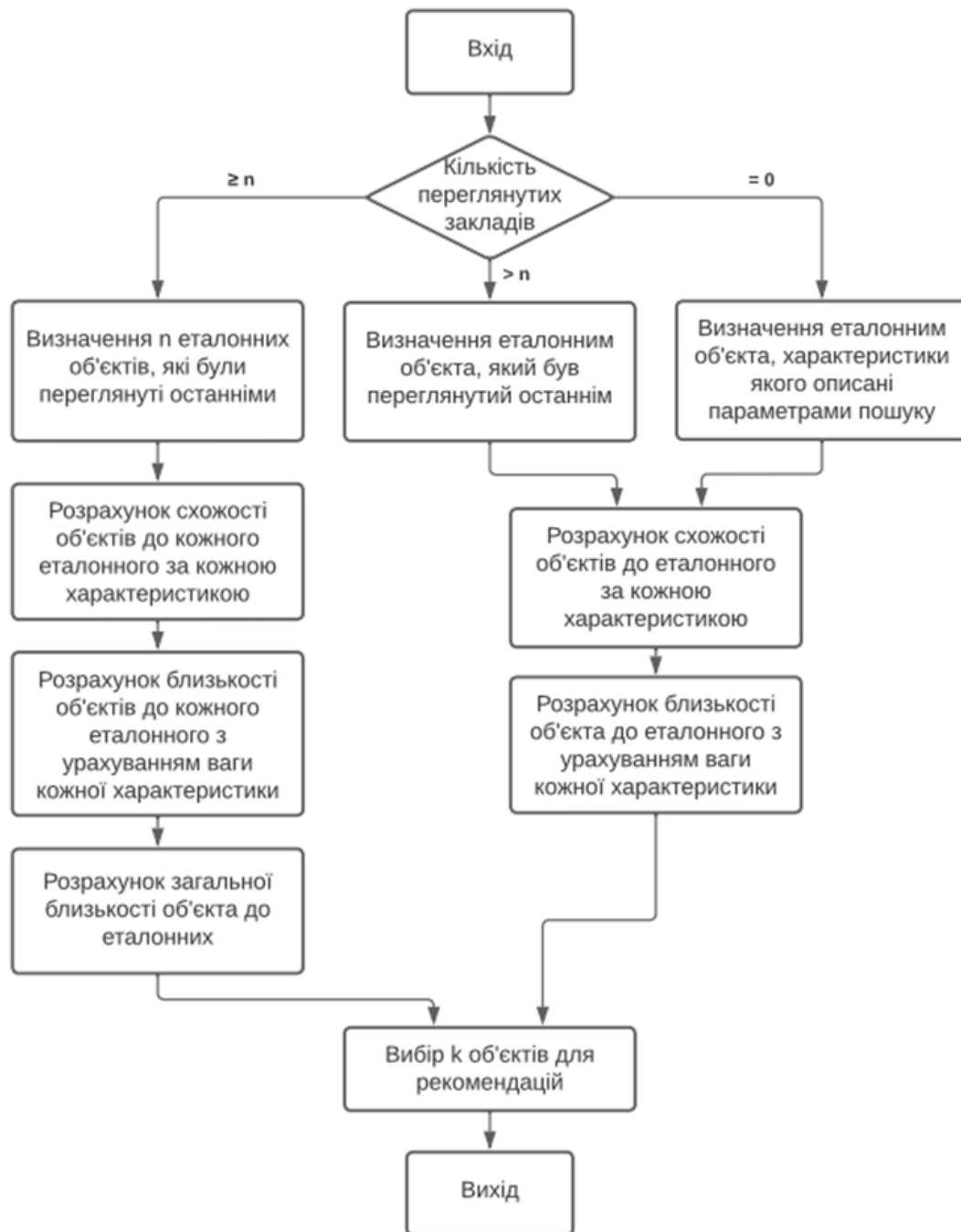


Рис.2 Блок-схема роботи алгоритму рекомендаційної системи

Спочатку алгоритм перевіряє кількість закладів, які переглянув користувач після застосування пошуку. У результаті перевірки виконується один із сценаріїв відповідно до варіантів:

- кількість переглянутих закладів є рівною нулю;
- кількість переглянутих закладів є меншою за встановлене порогове значення;
- кількість переглянутих закладів є більшою за встановлене порогове значення.

Якщо виконується перша умова, то еталонним вважається об'єкт, характеристиками якого є параметри фільтрації, а локація визначається місцезнаходженням користувача. Якщо виконується друга умова, то еталонним вважається останній переглянутий користувачем об'єкт. Для обох випадків за допомогою формул (3) і (5) обчислюються оцінки близькості інших об'єктів до вибраного центру, і вибирається  $k$  найбільш відповідних.

У випадку, коли кількість переглянутих закладів є більшою за встановлене порогове значення, останні  $n$  переглянутих оголошень вважаються «ідеальною» точкою. Потім за допомогою формул (3) і (5) обчислюються оцінки близькості інших об'єктів до кожного вибраних еталонних об'єктів. Після чого за допомогою формули (4) обчислюється остаточна оцінка близькості відносно вибраного центру, і вибирається  $k$  найбільш відповідних.

У експериментальній частині даної роботи проводиться пошук таких параметрів для розрахунків, які найкраще передбачають вподобання користувачів. Результати рекомендацій порівнюються евристично для визначення показників, які найкраще відповідають потребам користувачів. Підбір цих параметрів відбувається через аналіз правил, які діють у предметній області, а також через врахування уявлень середнього користувача. Важливо враховувати, які критерії важливі для такого користувача під час пошуку рекомендацій.

Розроблена СР для предметної області системи пошуку закладів сфери послуг використовує наступні характеристики (критерії) та їх важливість:

- сфера послуг;
- адміністративний район;
- послуга – 0.3;
- тип закладу – 0.05;
- вартість надання послуги – 0.2;
- відстань до еталонного закладу – 0.3;
- рейтинг закладу – 0.15.

Для розрахунку близькості характеристик за формулою 1.5 значення параметрів представлено в таблиці (1).

Таблиця 1. Значення параметрів розрахунку оцінки близькості

Назва критерію, одиниці вимірювання	Назва параметра		
	offset	scale	decay
Вартість, грн	5%	15%	0.5
Відстань, км	0.5	3	0.5
Рейтинг	0	1	0.5

Тестування розробленого алгоритму рекомендацій показало, що він здатен створювати релевантні рекомендації, що відповідають запитам користувачів. Алгоритм ефективно визначає бажане місцезнаходження закладу та враховує історію перекладів. Успішне тестування підкреслює готовність алгоритму до використання в реальних умовах.

#### 4 Висновки

У роботі проаналізовано реалізацію системи рекомендацій для складної предметної сфери та алгоритм її впровадження, розглянуто різні категорії систем рекомендацій та особливості алгоритмів їх побудови, області застосування, переваги та недоліки. Для складної предметної області системи пошуку закладів сфери послуг реалізовано та впроваджено систему рекомендацій гібридного типу.

Розроблений алгоритм об'єднує в собі підходи до розробки систем рекомендацій на основі знань та фільтрації на основі вмісту. Перевагами описаного алгоритму є відсутність необхідності зберігати та використовувати у розрахунках інформацію про попередні сеанси користувачів, вирішення проблеми холодного запуску та відсутності в базі даних системи об'єктів, що повністю відповідають вказаним параметрам пошуку, формування актуальних рекомендацій для поточного сеансу користувача та вирішення проблеми визначення місцезнаходження для рекомендованого закладу.

Запропонований підхід може бути використано для розробки та впровадження рекомендаційного алгоритму у складних предметних сферах, де дані про геолокацію важливі для надання рекомендацій.

#### ЛІТЕРАТУРА

1. I. MacKenzie, C. Meyer, S. Noble, How retailers can keep up with consumers. McKinsey & Company, 2013. URL: <https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers> (Last accessed: 29.08.2023).



2. C. C. Aggarwal, An Introduction to Recommender Systems. In *Recommender Systems*; Springer: Cham, Switzerland, 2016. URL: <https://www.springer.com/gp/book/9783642031205> (Last accessed: 30.08.2023).
3. Yang X., Guo Y., Liu Y., Steck H., A survey of collaborative filtering based social recommender systems. *Comput. Commun.* 2014, 41, pp. 1–10. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0140366413001722> (Last accessed: 30.08.2023).
4. Ding Z., Li X., Jiang C, Zhou M., Objectives and state-of-the-art of location-based social network recommender systems. *ACM Comput. Surv. CSUR* 2018, 51, pp. 1–28. URL: <https://dl.acm.org/doi/abs/10.1145/3154526> (Last accessed: 05.09.2023).
5. Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, R. Kashef, Recommendation Systems: Algorithms, Challengers, Metrics, and Business. *Appl. Sci.*, 2020. URL: <https://www.mdpi.com/2076-3417/10/21/7748#B10-applsci-10-07748> (Last accessed: 05.09.2023).
6. B.M. Sarwar, G. Karypis, J.A. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms. In *WWW '01, Proceedings of the 10th International Conference on World Wide Web*, Hong Kong, China, 1–5 May 2001; Association for Computing Machinery: New York, NY, USA, 2001; Volume 1, pp. 285–295. URL: <https://dl.acm.org/doi/abs/10.1145/371920.372071> (Last accessed: 05.09.2023).
7. J.B. Schafer, D. Frankowski, J. Herlocker, S. Sen, Collaborative filtering recommender systems. In *The Adaptive Web*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 291–324. URL: [https://link.springer.com/chapter/10.1007/978-3-540-72079-9\\_9](https://link.springer.com/chapter/10.1007/978-3-540-72079-9_9) (Last accessed: 07.09.2023).
8. Al-Shamri, M.Y.H., User profiling approaches for demographic recommender systems. *Knowl. Based Syst.* 2016, 100, pp. 175–187. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0950705116001192> (Last accessed: 10.09.2023).
9. Deng F., Utility-based recommender systems using implicit utility and genetic algorithm. In *Proceedings of the 2015 International Conference on Mechatronics, Electronic, Industrial and Control Engineering (MEIC-15)*, Shenyang, China, 1–3 April 2015; Atlantis Press: Amsterdam, The Netherlands, 2015. URL: <https://www.atlantispress.com/proceedings/meic-15/19830> (Last accessed: 11.09.2023).
10. R. Burke, Knowledge-based recommender systems. In *Encyclopedia of Library and Information Systems*; CRC Press: Boca Raton, FL, USA, 2000; Volume 69, (Suppl. 32), pp. 175–186. URL: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=dc133144d431fc3b75c8de27f6b21da6eb5bc1b> (Last accessed: 11.09.2023).
11. A.A. Kardan, M. Ebrahimi, A novel approach to hybrid recommendation systems based on association rules mining for content recommendation in asynchronous discussion groups. *Inf. Sci.* 2013, 219, pp. 93–110. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0020025512004756> (Last accessed: 18.09.2023).

#### REFERENCES

1. MacKenzie, C. Meyer, S. Noble, How retailers can keep up with consumers. McKinsey & Company, 2013. URL: <https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers> (Last accessed: 29.08.2023).
2. C. C. Aggarwal, An Introduction to Recommender Systems. In *Recommender Systems*; Springer: Cham, Switzerland, 2016. URL: <https://www.springer.com/gp/book/9783642031205> (Last accessed: 30.08.2023).
3. Yang X., Guo Y., Liu Y., Steck H., A survey of collaborative filtering based social recommender systems. *Comput. Commun.* 2014, 41, pp. 1–10. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0140366413001722> (Last accessed: 30.08.2023).
4. Ding Z., Li X., Jiang C, Zhou M., Objectives and state-of-the-art of location-based social network recommender systems. *ACM Comput. Surv. CSUR* 2018, 51, pp. 1–28. URL: <https://dl.acm.org/doi/abs/10.1145/3154526> (Last accessed: 05.09.2023).
5. Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, R. Kashef, Recommendation Systems: Algorithms, Challengers, Metrics, and Business. *Appl. Sci.*, 2020. URL: <https://www.mdpi.com/2076-3417/10/21/7748#B10-applsci-10-07748> (Last accessed: 05.09.2023).

6. B.M. Sarwar, G. Karypis, J.A. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms. In WWW '01, Proceedings of the 10th International Conference on World Wide Web, Hong Kong, China, 1–5 May 2001; Association for Computing Machinery: New York, NY, USA, 2001; Volume 1, pp. 285–295. URL: <https://dl.acm.org/doi/abs/10.1145/371920.372071> (Last accessed: 05.09.2023).
  7. J.B. Schafer, D. Frankowski, J. Herlocker, S. Sen, Collaborative filtering recommender systems. In The Adaptive Web; Springer: Berlin/Heidelberg, Germany, 2007; pp. 291–324. URL: [https://link.springer.com/chapter/10.1007/978-3-540-72079-9\\_9](https://link.springer.com/chapter/10.1007/978-3-540-72079-9_9) (Last accessed: 07.09.2023).
  8. Al-Shamri, M.Y.H., User profiling approaches for demographic recommender systems. Knowl. Based Syst. 2016, 100, pp. 175–187. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0950705116001192> (Last accessed: 10.09.2023).
  9. Deng F., Utility-based recommender systems using implicit utility and genetic algorithm. In Proceedings of the 2015 International Conference on Mechatronics, Electronic, Industrial and Control Engineering (MEIC-15), Shenyang, China, 1–3 April 2015; Atlantis Press: Amsterdam, The Netherlands, 2015. URL: <https://www.atlantispress.com/proceedings/meic-15/19830> (Last accessed: 11.09.2023).
  10. R. Burke, Knowledge-based recommender systems. In Encyclopedia of Library and Information Systems; CRC Press: Boca Raton, FL, USA, 2000; Volume 69, (Suppl. 32), pp. 175–186. URL: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=dc133144d431fc3b75c8de27f6b21da6eb5bc1b> (Last accessed: 11.09.2023).
- A.A. Kardan, M. Ebrahimi, A novel approach to hybrid recommendation systems based on association rules mining for content recommendation in asynchronous discussion groups. Inf. Sci. 2013, 219, pp. 93–110. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0020025512004756> (Last accessed: 18.09.2023).

## Development of a hybrid-based recommendation system

A.O. Bielan, L.V. Vasilieva

**Bielan Anzhelika**

*Student*

*V. N. Karazin Kharkiv National University  
6 Svobody Sq., Kharkiv, Ukraine, 61022*

**Vasilieva Larisa**

*Candidate of Biological Sciences, Associate Professor*

*V. N. Karazin Kharkiv National University  
6 Svobody Sq., Kharkiv, Ukraine, 61022*

Modern Internet and web services are flooded with a vast amount of information, which is becoming more challenging for users. Recommendation systems aim to address this information overload issue while personalizing the user experience by providing precise, personalized recommendations based on their preferences. The main goal of this work is implementing a recommendation system for a complex subject area and the algorithm of its integration into the sector of searching for the service providers. The article briefly outlines the main approaches and algorithms used in recommendation systems, highlighting their areas of application, advantages, and disadvantages, as well as description of the implementation of a recommendation system for a complex subject area and the algorithm for its integration into the sector of searching for the service providers. One distinctive feature of the system under development is the importance of geolocation data for the generation of recommendations. An algorithm for a hybrid recommendation system that combines knowledge-based and content-based filtering approaches has been developed. The advantages of the described algorithm include the absence of the need to store and use information about previous user sessions in the calculations, as well as addressing cold-start issues, generating real-time recommendations for the current user session, and solving the problem of determining the location for recommended establishments. The proposed approach can be used for developing and implementing recommendation algorithms in complex subject areas where geolocation data is crucial for providing recommendations.

**Keywords:** *recommendation system, recommendation algorithm, hybrid-based recommendation system, content-based filtering, knowledge-based recommenders*

**How to quote:** Bielan A., Vasilieva L., “Development of a hybrid-based recommendation system”, *Bulletin of V.N. Karazin Kharkiv National University, series “Mathematical modelling. Information technology. Automated control systems*, vol. 57, pp. 22-31, 2023.

<https://doi.org/10.26565/2304-6201-2023-57-02>