

С. Ф. Чалий, В. О. Лещинський, І. О. Лещинська

Харківський національний університет радіоелектроніки, Харків, Україна

ДОПОВНЕННЯ ВХІДНИХ ДАНИХ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ В СИТУАЦІЇ ЦИКЛІЧНОГО ХОЛОДНОГО СТАРТУ З ВИКОРИСТАННЯМ ТЕМПОРАЛЬНИХ ОБМЕЖЕНЬ ТИПУ «NEXT»

Предметом вивчення в статті є процеси формування рекомендованого списку товарів та послуг в ситуації циклічного холодного старту рекомендаційної системи. Така ситуація характеризується циклічною зміною інтересів користувачів, що потребує уточнення вхідних даних для побудови рекомендацій. **Метою** є розробка методу доповнення вхідних даних для побудови рекомендацій непостійним користувачам, що змінюють свої вимоги, з використанням темпоральних обмежень типу «Next». **Завдання:** виділити базові особливості темпоральних залежностей в рекомендаційних системах; розробити концепцію коригування вхідних даних на основі темпоральних обмежень типу «Next». Отримані наступні **результати.** Виконано структурування темпоральних залежностей з урахуванням особливості їх застосування в рекомендаційних системах. Показано, що для опису поведінки користувача в таких системах можна використовувати залежності типу «Next» між послідовним у часі вибором одного й того ж об'єкту, а також залежності типу «Future» з проміжним вибором інших товарів або послуг. Запропоновано концептуальний підхід до уточнення вхідних даних на основі темпоральних залежностей. В рамках представленого концептуального підходу запропоновано метод доповнення рекомендаційної системи на основі формування темпоральних обмежень типу «Next». **Висновки.** Наукова новизна отриманих результатів полягає в наступному. Запропоновано метод доповнення вхідних даних рекомендаційної системи в ситуації циклічного холодного старту з використанням темпоральних обмежень типу «Next». Метод містить етапи узагальнення вхідних даних, формування темпоральних обмежень типу «Next», а також доповнення вхідних даних згідно отриманих обмежень. Запропонований метод дозволяє підвищити ефективність побудови рекомендацій для непостійних користувачів на основі формування обмежень, що відображають обов'язкові зміни інтересів відомих користувачів.

Ключові слова: рекомендаційні системи, темпоральні обмеження, персоналізація рекомендацій, формування рекомендацій.

Вступ

Рекомендаційні системи призначені для підтримки вибору споживача із множини запропонованих товарів та послуг в Інтернет - магазинах, системах заказу туристичних турів та готелів, системах продажу послуг з перегляду відео, тощо [1 – 3]. Алгоритми роботи рекомендаційних систем дають можливість прогнозувати набір товарів або послуг, що відповідають потребам конкретного споживача. Рекомендаційні системи при формуванні пропозицій користувачеві використовують дані про вибір схожих споживачів, а також про схожість характеристик запропонованих предметів та тих, що були вибрані споживачем раніше [4, 5].

Однією з ключових проблем, що виникає при побудові та удосконаленні рекомендаційних систем, є проблема холодного старту. Остання виникає в ситуації неповноти даних про вибір товарів та послуг новими споживачами [6]. При появі нового користувача рекомендаційної системи або нового товару чи послуги вони розглядаються як «холодні», тому що рекомендаційна система не має достатньо інформації для формування релевантних пропозицій новим користувачам або рекомендацій щодо нових товарів та послуг.

Для формування рекомендацій необхідно, щоб користувач вибрав деякий набір товарів та послуг за власною ініціативою.

Після вибору користувачі перестають розглядатись «холодні» і тому рекомендації для них мо-

жуть бути сформовані за допомогою відомих алгоритмів.

Проблема холодного старту виникає не лише для нових, але й для нерегулярних користувачів. Останні заходять на відповідний сайт електронної комерції через великі інтервали часу. Така поведінка зазвичай характерна для споживачів, що купують дорогу побутову техніку, туристичні подорожі, тощо [7].

Фактично дана ситуація є проміжною між традиційним холодним стартом та поведінкою постійного споживача. Знання про залежності між вибором споживачів в даному випадку є неявними і вони не в повному обсязі відображені в рекомендаційній системі, що значно утруднює формування персоналізованих рекомендацій.

Традиційно при вирішенні проблеми холодного старту виконується аналіз контексту [8] вхідних даних, після чого виконується колаборативна фільтрація [9].

При аналізі контексту враховуються також характеристики поведінки користувачів. Останні можуть бути побудовані з використанням градієнтного спуску [10], та представлені у вигляді темпорального графу [11] або нейромережевої моделі [12]. Моделювання поведінки користувача передбачає використання повного масиву вхідних даних, що підвищує обчислювальні витрати. Однак такі підходи мають досить високу точність.

Альтернативні підходи пов'язані із фільтрацією вхідних даних, під час якої видаляється нереле-

вантна для поточної рекомендації інформація. Таке коригування масиву вхідних даних виконується на основі багатопланового графового представлення поведінки користувача [13], а також на основі самонавчання [14] та активного навчання [15].

Ключова перевага останньої групи підходів полягає в тому, що вони є доповненням до традиційних методів побудови рекомендацій у випадку вирішення проблеми холодного старту. Однак в цілому існуючі підходи до вирішення проблеми холодного старту не враховують притаманне циклічному холодному старту поєднання змін у інтересах споживачів та неповноту інформації про вибір цих споживачів, що і свідчить про актуальність даної проблеми.

Слід зазначити, що моделі таких неявних залежностей у їх темпоральному аспекті можуть бути побудовані шляхом аналізу послідовностей вибору користувачів [16].

Відповідно, для вирішення розглянутої проблеми доцільно доповнити вхідні дані нерегулярних користувачів з використанням сформованих темпоральних обмежень циклу вибору товарів та послуг для уже відомих споживачів. Це дасть можливість в подальшому використовувати традиційні методи для формування пропозицій в ситуації циклічного холодного старту.

Метою статті є розробка методу доповнення вхідних даних для побудови рекомендацій непостійним користувачам, що змінюють свої вимоги, з використанням темпоральних обмежень.

Даний метод має забезпечити можливість побудови рекомендацій для нерегулярних користувачів, інтереси яких змінюються з часом.

Для досягнення поставленої мети потрібно вирішити такі задачі:

- виділити базові особливості темпоральних залежностей в рекомендаційних системах;
- розробити концепцію коригування вхідних даних у ситуації циклічного холодного старту;
- розробити метод доповнення вхідних даних на основі темпоральних обмежень типу «Next».

Структуризація особливостей темпоральних залежностей

Темпоральні залежності згідно результатів робіт [17, 18] представляються темпоральними правилами, що пов'язують пари подій з однієї послідовності подій. Адаптація цих темпоральних залежностей для задачі побудови рекомендацій дає можливість визначити темпоральні правила як зв'язки у часі між подіями вибору визначеної підмножини товарів та послуг для заданої підмножини користувачів. Такі залежності визначають типові зміни поведінки споживачів з часом і тому можуть бути використані для визначення обмежень у поведінці нерегулярних користувачів рекомендаційної системи.

Побудова рекомендацій в цілому направлена на звуження можливостей вибору користувача таким чином, щоб спростити його вибір, найкраще задовольнивши його інтереси. Такі інтереси пред-

ставлені відомою множиною вибраних цим користувачем товарів та послуг. У випадку циклічного холодного старту минулий вибір користувача вже не є актуальним. Однак він може коригуватись із відомим вибором інших користувачів на даному інтервалі часу, наприклад, із сезонним вибором товарів, подорожей, вибором популярних в даний момент співаків, фільмів, тощо.

Тому задача побудови рекомендацій в даній ситуації направлена на звуження можливостей вибору нерегулярного користувача на основі темпоральних обмежень, які задаються вибором інших користувачів.

Для вирішення цієї задачі пропонується доповнити вхідні дані для нерегулярного користувача записами обмежень по вибраній підмножині користувачів на заданому інтервалі часу.

Такі темпоральні обмеження можуть бути задані за допомогою темпоральних операторів X (Next) та F (Future) [17]. Перший з них задає зв'язок між двома послідовними виборами споживача.

Такий зв'язок з урахуванням особливостей використання рекомендаційних систем може бути інтерпретований в двох аспектах:

- послідовний у часі вибір одного й того ж об'єкту (товару, послуги) різними споживачами;
- послідовний вибір одним споживачем (або визначеною групою схожих споживачів) різних товарів та послуг.

В першому випадку темпоральне правило типу «Next» визначає послідовність зміни інтересу споживачів до визначеної групи товарів та послуг. Така зміна інтересів задається на парі послідовних інтервалів часу.

В другому випадку дане темпоральне правило визначає послідовність зміни інтересів вибраної групи споживачів. Зазначена група може бути відібрана за рядом ознак (вік, місце проживання, освіта, тощо).

Темпоральні правила типу Future відрізняються тим, що вони задають зв'язок між довільною парою виборів користувачів, між якими було здійснено інший вибір. Тобто такі правила визначають зміну інтересів користувачів для значних інтервалів часу.

Комбінація вказаних правил дозволяє моделювати динамічний аспект інтересів споживачів.

Концепція коригування вхідних даних у ситуації циклічного холодного старту

Ключова ідея запропонованого підходу полягає у використанні темпоральних моделей поведінки відомих споживачів для доповнення вхідних даних нерегулярних користувачів рекомендаційної системи. Це дає можливість побудувати перелік пропозицій товарів та послуг у разі відсутності або неактуальності вхідних даних щодо вибору поточного користувача рекомендаційної системи у разі зміни його інтересів протягом тривалого періоду часу.

Узагальнену схему запропонованої концепції представлено на рис. 1.

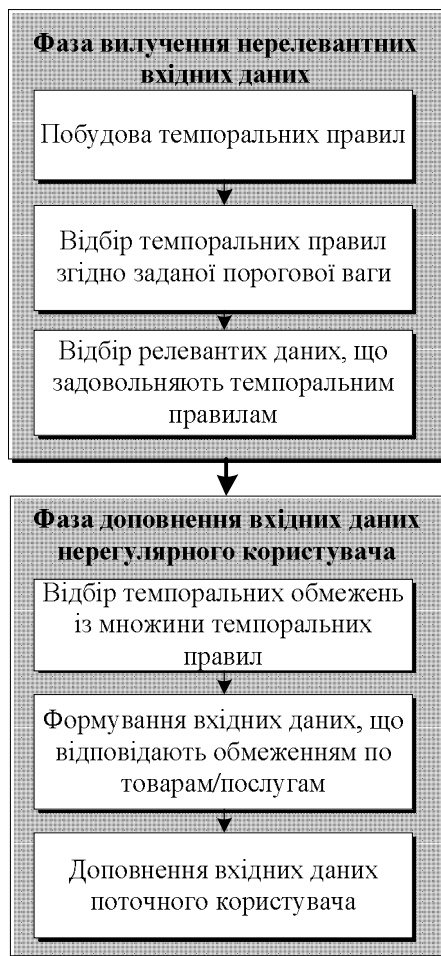


Рис. 1. Концептуальна схема коригування даних у ситуації циклічного холодного старту

Представлена на даному рисунку послідовність містить такі дві ключові фази:

- видалення нерелевантної інформації;
- доповнення інформації на основі темпоральних обмежень.

Видалення нерелевантної інформації дає можливість не лише підвищити точність рекомендацій, але й понизити обчислювальні витрати на їх побудову. Останнє є особливо важливим у випадку формування рекомендованого переліку товарів та послуг у режимі он-лайн.

Доповнення вхідних даних задає «коридор вибору» для користувача. Як уже було зазначено вище, такий «коридор вибору» визначається допустимою поведінкою (в сенсі послідовності виборів) споживачів.

Поведінка споживача може бути представлена темпоральними правилами та темпоральними обмеженнями. Відмінність обмежень від темпоральних правил полягає в тому, що вони виконуються на всій визначеній послідовності інтервалів часу.

Таким чином, згідно запропонованої концепції при реалізації обох розглянутих фаз виконується побудова темпоральних залежностей та їх подальше використання для коригування вхідних даних.

На першій фазі формуються темпоральні правила, які дають можливість із заданим рівнем ймовірності відібрати вхідні дані.

На другій фазі формуються темпоральні обмеження, які задають обов'язкові зміни поведінки для визначеного інтервалу часу. Ці обмеження визначають мінімальний набір покупок, який був виконаний споживачами. Тому вказаний набір використовується в якості вхідних даних для нерегулярного споживача.

Метод доповнення вхідних даних рекомендаційної системи в ситуації циклічного холодного старту

Вхідними даними методу є:

- журнал (лог) продажів товарів та послуг;
- інтервал часу $[\tau_1, \tau_N]$ для відбору даних, що відповідають актуальним змінам поведінки користувачів;
- рівень деталізації часу, тобто рівень базових інтервалів часу $\Delta\tau$, відносно яких формуються темпоральні залежності (година, день, тиждень, тощо).

Журнал продажів має наступну структуру:

$$L = \{l_m\}, \quad (1)$$

$$l_m = (i_j, u_k, \tau_m, p_m),$$

де u_k – користувач рекомендаційної системи;

i_j – об'єкт (товар або послуга), який вибрав користувач;

τ_m – момент часу вибору користувача;

p_m – кількість одиниць вибраних об'єктів.

Метод містить у собі такі етапи.

Етап 1. Відбір підмножини записів із журналу L для заданого інтервалу часу:

$$L^* = \{l_m : \tau_m \in [\tau_1, \tau_N]\}, \quad (2)$$

$$L^* \subseteq L.$$

Етап 2. Поєднання покупок l_m для заданого рівня деталізації часу $\Delta\tau$ та заданого об'єкту i_j у множини l_r^{**} :

$$L^{**} = \{l_r^{**}\}, \quad (3)$$

$$l_r^{**} = (p_r, i_j, \Delta\tau_r) : \tau_m \in \Delta\tau_r,$$

де p_r – кількість вибраних (куплених) одиниць товару на заданому s – інтервалі часу (для заданого рівня деталізації) $\Delta\tau_r$.

Значення p_r з урахуванням умови $\tau_m \in \Delta\tau_r$ у виразі (3) обраховується традиційно:

$$p_r = \sum_{i_j \in l_m} p_m. \quad (4)$$

Еман 3. Формування обмежень

$$C^{Next} = \{c_{j,r}^{Next}\}$$

типу «Next» для пар послідовних інтервалів $\Delta\tau_r$:

$$c_{j,r}^{Next} = l_r^{**} X l_{r+1}^{**} : i_j \in l_r^{**}, i_j \in l_{r+1}^{**} \Rightarrow \exists c_{j,r}^{Next}, \quad (5)$$

де $c_{j,r}^{Next}$ – обмеження по вибору об'єкту i_j на послідовних інтервалах часу $\Delta\tau_r$, $\Delta\tau_{r+1}$.

Еман 4. Доповнення вхідних даних L для неперіодичного користувача парами l_r^{**}, l_{r+1}^{**} , що відображають обов'язкові для інтервалу $[\tau_1, \tau_N]$ зміни вподобань споживачів.

Відзначимо, що у загальному випадку значення p_r для цих нових записів вхідних даних може бути встановлено рівним 1, щоб мінімізувати вплив поєднання покупок на етапі 2 методу.

Висновки

Розглянуто проблему побудови рекомендацій неперіодичним користувачам у випадку циклічного холодного старту.

Останній характеризується циклічною зміною інтересів таких користувачів, що потребує уточ-

нення вхідних даних. Для вирішення цієї проблеми запропоновано реалізувати уточнення вхідних даних з урахуванням темпоральних аспектів вибору вже відомих користувачів.

Виконано структурування темпоральних залежностей.

Показано, що для опису поведінки користувача в рекомендаційних системах можна використовувати залежності між послідовним у часі вибором одного й того ж об'єкту, а також вибором на інтервалах часу, між якими є проміжний вибір інших товарів або послуг.

Запропоновано концептуальний підхід до уточнення вхідних даних на основі темпоральних залежностей.

На основі представленого концептуального підходу запропоновано метод доповнення вхідних даних рекомендаційної системи в ситуації циклічного холодного старту з використанням темпоральних обмежень типу «Next».

Метод містить етапи узагальнення вхідних даних, формування темпоральних обмежень типу «Next», а також доповнення вхідних даних згідно отриманих обмежень.

Запропонований метод дозволяє підвищити ефективність формування рекомендацій для неперіодичних користувачів з урахуванням змін інтересів відомих споживачів.

REFERENCES

- Linden G., Smith B. and York J. (2003), "Amazon.com recommendations: Item-to-item collaborative filtering", *Internet Computing*, IEEE 7, 1, pp. 76–80.
- Jannach D., Gedikli F., Karakaya Z., Juwig O. (2012) Recommending Hotels based on Multi-Dimensional Customer Ratings. In: Fuchs M., Ricci F., Cantoni L. (eds) *Information and Communication Technologies in Tourism 2012*. Springer, Vienna, pp 320-331.
- Bennet J. and Lanning S. (2007) "The Netflix Prize", *Proceedings of KDD cup and workshop*, available at: <http://www.netflixprize.com> (last accessed May 31, 2019).
- Aggarwal C. (2017). *Recommender Systems: The Textbook*, New York: Springer. 498 p.
- Adomavicius G. and Tuzhilin A. (2005), "Towards the Next Generation of Recommender Systems" A Survey of the State-of-the-Art and Possible Extensions, *IEEE Transactions on Knowledge and Data Engineering*, No. 17, pp. 634–749.
- Son Le Hoang (2016). Dealing with the new user cold-start problem in recommender systems: A comparative review. *Information Systems*, 58, 87-104.
- Bernardi L., Kamps J., Kiseleva J, Müller M. (2015). The Continuous Cold Start Problem in e-Commerce Recommender Systems. *CoRR abs/1508.01177*.
- Чалий С.Ф., Лещинський В.О., Лещинська І.О. (2018). Моделювання контексту в рекомендаційних системах. *Проблеми інформаційних технологій*, 1(023), 21-26.
- Braunhofer M. (2014). Hybrid Solution of the Cold-Start Problem in Context-Aware Recommender Systems. *UMAP 2014: User Modeling, Adaptation, and Personalization*, 484-489.
- Koren, Y. (2009). Collaborative Filtering with Temporal Dynamics. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 447-456.
- Xiang L., Yuan Q. (2010). Temporal Recommendation on Graphs via Long-and Short-term Preference Fusion. *KDD'10 of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 723-732.
- Elahi M., Ricci F., Rubens N. (2016). A survey of active learning in collaborative filtering recommender systems. *Computer Science Review*, 20, 29-50.
- Chalyi S., Pribylnova I. (2019). The method of constructing recommendations online on the temporal dynamics of user interests using multilayer graph. *EUREKA: Physics and Engineering*, 3, 13-19.
- Luo C., Cai X. (2014). Self-training Temporal Dynamics Collaborative Filtering. *PAKDD'14*, 461-472.
- Zhu Y., Lin J., He S., Wang B., Guan Z., Liu H., and Cai D. (2018). Addressing the item cold-start problem by attribute-driven active learning." *arXiv preprint arXiv:1805.09023*.
- Kalynychenko O., Chalyi S., Bodyanskiy Y., Golian V., Golian N. (2013). Implementation of search mechanism for implicit dependences in process mining. *2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced*

- Computing Systems (IDAACS). Institute of Electrical and Electronics Engineers (IEEE). DOI: 10.1109/IDAACS.2013.6662657
17. Levykin V., Chala O. (2018). Development of a method of probabilistic inference of sequences of business process activities to support business process management. Eastern-European Journal of Enterprise Technologies, 5/3(95), 16-24. DOI: 10.15587/1729-4061.2018.142664.
18. Levykin V., Chala O. (2018). Method of determining weights of temporal rules in markov logic network for building knowledge base in information control system. EUREKA: Physics and Engineering, 5(18), 3-10.

Рецензент: д-р техн. наук, проф. Г. А. Кучук,
 Національний технічний університет «ХПІ», Харків
 Received (Надійшла) 27.06.2019
 Accepted for publication (Прийнята до друку) 21.08.2019

**Дополнение входных данных рекомендательной системы
 в ситуации циклического холодного старта
 с использованием темпоральных ограничений типа «Next»**

С. Ф. Чалый, В. А. Лещинский, И. А. Лещинская

Предметом изучения в статье являются процессы формирования рекомендованного списка товаров и услуг в ситуации циклического холодного старта рекомендательной системы. Такая ситуация характеризуется циклической сменой интересов пользователей, что требует уточнения исходных данных при построении рекомендаций. **Целью** является разработка метода дополнения входных данных для построения рекомендаций непостоянным пользователям, которые меняют свои требования, с использованием темпоральных ограничений типа «Next». **Задачи:** выделить базовые особенности темпоральных зависимостей в рекомендательных системах; разработать концепцию корректировки входных данных в ситуации циклического холодного старта; разработать метод дополнения входных данных на основе темпоральных ограничений типа «Next». Получены следующие **результаты**. Выполнена структуризация темпоральных зависимостей с учетом особенности их применения в рекомендательных системах. Показано, что для описания поведения пользователя в таких системах можно использовать зависимости типа «Next» между последовательным во времени выбором одного и того же объекта, а также зависимости типа «Future» с промежуточным выбором других товаров или услуг. Предложен концептуальный подход к уточнению входных данных на основе темпоральных зависимостей. В рамках представленного концептуального подхода предложен метод дополнения входных данных рекомендательной системы на основе формирования темпоральных ограничений типа «Next». **Выводы.** Научная новизна полученных результатов заключается в следующем. Предложен метод дополнения входных данных рекомендательной системы в ситуации циклического холодного старта с использованием темпоральных ограничений типа «Next». Метод включает этапы обобщения исходных данных, формирования темпоральных ограничений типа «Next», а также дополнения входных данных согласно полученным ограничениям. Предложенный метод позволяет повысить эффективность построения рекомендаций для непостоянных пользователей на основе формирования ограничений, отражающие обязательные изменения интересов известных пользователей.

Ключевые слова: рекомендательные системы, темпоральные ограничения, персонализация рекомендаций, формирование рекомендаций.

**Supplementing the input of the recommendation system
 in a cyclic cold start situation
 using temporal constraints of the “next” type**

S. Chalyi, V. Leshchynskiy, I. Leshchynska

The **subject matter** of the article is the formation of the recommended list of goods and services in a situation of a cyclical cold start of the recommendation system. This situation is characterized by a cyclical change in the interests of users, which requires clarification of the source data when building recommendations. The goal is to develop a method of supplementing the input data to build recommendations for non-permanent users who change their requirements, using temporal constraints such as “Next”. **Tasks:** to highlight the basic features of temporal dependencies in recommendation systems; to develop the concept of adding input data in a situation of cyclic cold start; to develop a method for supplementing input data based on temporal constraints of the “Next” type. The following **results** are obtained. The temporal dependencies are structured taking into account the features of their application in recommender systems. It is shown that to describe user behavior in such systems, you can use dependencies of the “Next” type between time-consistent selection of the same object, as well as dependencies of the “Future” type with an intermediate choice of other goods or services. A conceptual approach to refine the input data based on temporal dependencies is proposed. Within the framework of the presented conceptual approach, a method is proposed for supplementing the input data of the recommendation system based on the formation of temporal constraints of the “Next” type. **Conclusions.** The scientific novelty of the results is as follows. A method is proposed for supplementing the input data of a recommender system in a cyclic cold start situation using temporal constraints of the “Next” type. The method includes the steps of summarizing the initial data, forming temporal constraints of the “Next” type, and also supplementing the input data according to the obtained constraints. The proposed method allows to increase the efficiency of constructing recommendations for non-regular users based on the formation of restrictions that reflect mandatory changes in the interests of known users.

Keywords: recommendation systems, temporal restrictions, personalization of recommendations, formation of recommendations.