

## МЕТОД ОБ'ЄДНАННЯ ОЦІНОК ДЛЯ РЕКОМЕНДАЦІЙНИХ СИСТЕМ НА ОСНОВІ КОЛАБОРАТИВНОЇ ФІЛЬТРАЦІЇ

Савчук Тамара, Сакалюк Антон

Вінницький національний технічний університет

### Анотація

Забезпечення точних рекомендацій має важливе значення для систем електронної комерції, щоб допомогти користувачам у прийнятті ефективних рішень вибору з множини варіантів. Колаборативна фільтрація є широко вживаним методом для генерації рекомендацій, основаних на оцінках користувачів, що мають спільні вподобання. Тим не менш, вона страждає від декількох притаманних їй проблем, як розрідженість даних і холодний старт. Для вирішення цих проблем був запропонований новий метод об'єднання оцінок, що включає інформацію соціальної довіри (тобто, довірених сусідів явно чи неявно зазначеної користувачами) у наданні рекомендації.

### Abstract

Providing high quality recommendations is important for e-commerce systems to assist users in making effective selection decisions from a plethora of choices. Collaborative filtering is a widely accepted technique to generate recommendations based on the ratings of like-minded users. However, it suffers from several inherent issues such as data sparsity and cold start. To address these problems, a novel method called "Merge ratings" was proposed to incorporate social trust information (i.e., trusted neighbors explicitly or implicitly specified by users) in providing recommendations.

### Вступ

Колаборативна фільтрація є одним з найвідоміших і широко використовуваних методів для генерації рекомендацій [1]. Даний евристичний підхід базується на тому, що оцінки тих, хто мають схожі вподобання, будуть також схожими. Тим не менш, для колаборативної фільтрації притаманні такі проблеми, як розрідженість даних і проблема холодного старту. Перша проблема полягає у складності пошуку достатньо надійних користувачів, що мають подібні вподобання, через те, що користувачі самостійно оцінюють тільки невелику частину об'єктів. Друга ж проблема відноситься до дилеми, що неможливо зробити точні висновки по рекомендаціям, поки користувач оцінив замало об'єктів.

Щоб вирішити ці проблеми, зазвичай використовується додаткова інформація з інших джерел про користувачів (наприклад, соціальні мережі), де існує встановлена довіра між користувачами. У даній роботі, довіра визначається як переконання в забезпеченні точної оцінки щодо вподобань активного користувача по відношенню до інших. Обидва види - неявна довіра (наприклад, [2]) і явна довіра (наприклад, [1, 3]) були досліджені в літературі. Неявна довіра виводиться з поведінки користувачів, наприклад, рейтинги, тоді як явна довіра визначається безпосередньо між користувачами. За визначенням, явна довіра є більш точною і надійною, ніж неявна.

### Метод об'єднання оцінок

У даній роботі запропоновано новий довірчий підхід, так зване «об'єднання оцінок довірених користувачів» шляхом включення довірених сусідів до явно довірених для кожного користувача в системі, з метою підвищити точність рекомендацій та подолати проблеми розрідженості даних і проблем холодного старту. Зокрема, було об'єднано рейтинги довірених сусідів користувача усередненням рейтингів від довірених сусідів, схожих на активного користувача. Якість об'єданого рейтингу вимірюється довірою враховуючи кількість рейтингів і співвідношення кількості конфліктів між позитивними і

негативними рейтингами. Набір об'єднаних рейтингів потім використовується для виявлення вподобань активного користувача та для знаходження довірених користувачів на основі подібності. Крім того, рейтинг довіри також враховується при розрахунку схожості користувачів.

Метод досягає високої точності рекомендацій, особливо ефективний для холодних користувачів порівняно з іншими аналогами.

Кожен користувач може оцінити певну кількість об'єктів, надаючи їм цілочисельну оцінку, як показано в табл. 1.

Таблиця 1 – Матриця оцінок

	i <sub>1</sub>	i <sub>2</sub>	i <sub>3</sub>	i <sub>4</sub>	i <sub>5</sub>	i <sub>6</sub>	i <sub>7</sub>	i <sub>8</sub>	i <sub>9</sub>
u <sub>1</sub>			5		?				
u <sub>2</sub>	5		4		3			2	
u <sub>3</sub>		4		3				1	
u <sub>4</sub>	3		5		2				
u <sub>5</sub>		4	4		3			3	
u <sub>6</sub>		3	3	5	5				
u <sub>7</sub>							5		4
u <sub>8</sub>			4		2			1	
u <sub>9</sub>			4		5			5	

Довіра може бути подана у вигляді матриці довіри (рис. 1). Якщо користувачі А довіряє користувачам В, а В – С, це вказує на те, що користувачі А довіряють користувачам С.

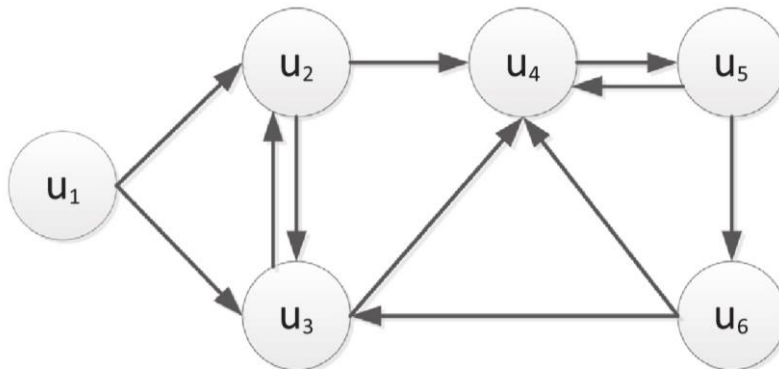


Рисунок 1 – Мережа довіри для користувача u<sub>1</sub>

Першим кроком методу об'єднання оцінок є визначення довірених сусідів для кожного користувача через «розповсюдження» довіри в мережі (рис. 1). Оцінка довіри між активним користувачем та іншими користувачами розраховується за формулою:

$$sim_{i,j} = \frac{\sum_{i \in r_{i \Delta j}} (r_{it} - \bar{r}_i)(r_{jt} - \bar{r}_j)}{\sqrt{\sum_{i \in r_{i \Delta j}} (r_{it} - \bar{r}_i)^2} \sqrt{\sum_{i \in r_{i \Delta j}} (r_{jt} - \bar{r}_j)^2}}$$

де  $r_{ij}$  – відстань між користувачами  $i$  та  $j$ .

Далі потрібно визначити найбільш довіреного сусіда:

$$csim_{i,j} = \frac{\sum_{k=1}^N r_{ik} \cdot r_{jk}}{\sqrt{\sum_{k=1}^N r_{ik}^2} \cdot \sqrt{\sum_{k=1}^N r_{jk}^2}}$$

Тоді коефіцієнт довіреності для сусідів активного користувача можна розрахувати за формулою:

$$P_{ui} = \bar{r}_i + \frac{\sum_{m=1}^c (r_{mi} - \bar{r}_m) \cdot c_{sim_{u,m}}}{\sum_{m=1}^c c_{sim_{u,m}}}$$

Тільки рейтинги довірених сусідів по об'єктах, що не оцінені активним користувачем, будуть об'єднані. Простіше кажучи, при виробленні рекомендацій, в активі користувача будуть знаходитись всі його власні оцінки, і оцінки довірених сусідів, так що в результаті буде сформовано новий більш точний рейтинг вподобань користувача.

## Висновки

Метод об'єднання оцінок має дві явних переваги порівняно з іншими методами. По-перше, він може ефективно поліпшити розрідженість даних і вирішити проблему холодного старту. Істотним викликом з двох проблем є те, що невелика кількість спільно оцінених об'єктів двома користувачами робить неможливим точно обчислити схожість користувачів, і, отже, знайти надійних подібних користувачів. Ще гірше, коли два користувача можуть не мати спільно оцінених об'єктів. Метод об'єднання вирішує проблему шляхом об'єднання рейтингів надійних сусідів з метою формування нового рейтингу, який використовується для представлення уподобань активного користувача.

По-друге, метод об'єднання може функціонувати в разі розрідженості рейтингів або розрідженості інформації довіри. Раніше багато цільових підходів, таких як MoleTrust [3] і TidalTrust [4] ставили рейтинги об'єктам тільки спираючись на рейтинг надаваних довірених сусідів. Тому, ці підходи страждали від проблем холодного старту, коли деякі користувачі можуть вказати тільки невелику кількість інших користувачів, їх довірених сусідів, тому ця проблема є загальною для багатьох інтернет-систем, особливо коли у користувачів відсутній стимул для об'єднування одного з одним. В цьому випадку точність буде обмежена, так як тільки кілька сусідів можуть бути включені в рекомендацію. На противагу цьому, метод об'єднання вирішує цю проблему. Зокрема, коли активний користувач не вказав ніяких надійних сусідів, але надав рейтинг певної кількості об'єктів, об'єднаний рейтинг буде точно такий же, як власний. Тобто метод об'єднання не матиме розбіжностей з традиційним методом колаборативної фільтрації. З іншого боку, коли користувач не оцінював об'єкти, але має декілька довірених сусідів, то рейтинги цих довірених сусідів можуть бути об'єднані. У кожному разі, метод є компетентним, щоб сформувати новий рейтинг, а отже, пом'якшити проблему холодного стару. Хоча метод не буде працювати, коли немає ні оцінок об'єктів, ні інформації про довірених сусідів активних користувачів. В цьому випадку буде необхідним інший вид інформації, щоб допомогти визначити вподобання користувачів.

## Список використаних джерел:

1. G. Guo, J. Zhang, D. Thalmann, A simple but effective method to incorporate trusted neighbors in recommender systems, in: Proceedings of the 20th International Conference on User Modeling, Adaptation and Personalization, 2012, pp. 114-125.
2. C. Haydar, A. Boyer, A. Roussanly, et al., Hybridising collaborative filtering and trust-aware recommender systems, in: Proceedings of the 8th International Conference on Web Information Systems and Technologies, 2012.
3. C.S. Hwang, Y.P. Chen, Using trust in collaborative filtering recommendation, in: New Trends in Applied Artificial Intelligence, 2007, pp. 1052-1060.
4. B. Knijnenburg, J. O'Donovan, S. Bostandjiev, A. Kobsa, Inspectability and control in social recommenders, in: Proceedings of the 6th ACM Conference on Recommender Systems, 2012.