Business Intelligence Management Studies
Vol. 12, No.46, Winter 2024; P 327-360
ims.atu.ac.ir
DOI: 10.22054/ims.2023.70722.2247

ATU PRESS

Original Research
Accepted: 03 December 2023
Revised: 04 October 2023
Received: 25 October 2022
eISSN: 2821-0816
ISSN: 2821-0964

Recommender Systems Based On Time and Trust Using Graph Based Community Detection

Fatemeh Rezaimehr
Ph.D. Student, Computer Engineering Faculty, Artificial Intelligence Department, K.N.Toosi University of technology, Tehran, Iran.

Chitra Dadkhah *
Assistant Professor, Computer Engineering Faculty, Artificial Intelligence Department, K.N.Toosi university of Technology, Tehran, Iran.

Abstract

Recently, the Internet has played a significant and substantial role in people's lives. However, the content available in the global web environment should align with users' daily needs, providing them with useful and up-to-date information tailored to their tastes. In this context, recommender systems assist users by suggesting items that closely match their preferences in less time. Today, with the exponential growth of data, the utilization of recommender systems has surged. Conversely, these systems encounter challenges such as evolving user preferences over time, cold start problem, sparsity within the user-item matrix, the infiltration of fake users in the systems, and their adverse impact on the recommendation lists. The objective of this paper is to propose a recommender system grounded in time and trust factors to enhance the efficiency and precision of system recommendations. Initially, the proposed system addresses the data sparsity dilemma by incorporating reliable implicit ratings into the user-item matrix. Subsequently, it constructs a weighted user-user network based on user rating timestamps and trust relationships among users, thereby mitigating the cold start problem and accounting for changing user preferences over time. The proposed recommender system employs a novel community detection algorithm introduced in this paper to identify the nearest neighbors of active users and recommends the top @k items based on the collaborative filtering

* Corresponding Author: dadkhah@ kntu.ac.ir
How to Cite: Rezaimehr, F., Dadkhah, Ch. (2024). The Impact of Gamification Elements on Customer Engagement, *Journal of Business Intelligence Management Studies*, 12(46), 327-360.

approach. Evaluation results of the proposed system, tested on a film recommender system using the Epinions dataset, demonstrate its superior efficiency compared to basic systems.

1. Introduction

Today, with the increasing tendency of users to use websites for obtaining information, online shopping, and using social networks for expressing personal opinions, the ways of obtaining information and establishing connections among users have undergone significant changes. Consequently, users are confronted with the big of data. Managing this data and selecting the appropriate options from this vast collection and presenting it to users is one of the main reasons for the development of information retrieval systems and search engines. In this regard, Recommendation Systems (RSs) help users choose the best options and recommend items that are closer to their preferences in the shortest possible time. Different models of RS such as collaborative filtering, content-based, knowledge-based, and newly developed context-aware RS, have been presented by researchers (Casillo et al., 2022). Each has its own advantages and disadvantages, which can be combined to create a hybrid RS. It should be noted that RS face challenges, including changes in user preferences over time, cold start for new users or items, sparsity of the user-item matrix, attack by fake users, and their negative impact on the recommendation list. In this paper, a time- and trust-based recommendation system is presented to enhance the performance and accuracy of recommendations. Our proposed system initially solves the data sparsity problem by adding reliable implicit ratings to the user-item rating matrix. It then generates a weighted user-user network based on the time of user feedback on items and trust relationships among users. This approach addresses the cold start problem and the change in user preferences over time. Our system is based on a novel community detection algorithm presented in this article, which identifies the nearest neighboring users with similar tastes to the active user and recommends the top-k items using the collaborative filtering method. The evaluation of the proposed system is performed on an Epinions dataset for a movie recommendation system. The evaluation uses metrics such as accuracy, recall, F1 score, mean absolute error, and root mean square error. The experimental results indicate the

superior performance of the proposed system compared to similar systems.

2. Literature Review

In the recent years, the researchers attempt to improve the accuracy of their recommendation for retaining the users and increasing the profit. Some of the papers has worked on optimizing the performance of their proposed RS using evolutionary algorithms (Tohidi & Dadkhah, 2020) and the others used the additional information such as time, location, etc. Trust-based RSs have been recently introduced to the community of computer science. Recent studies have shown that incorporating social factors or trust statements in RSs leads to the improvement of recommendation quality (P. Moradi & Ahmadian, 2015; S. Ahmadian, M. Meghdadi, & Afsharchi, 2018b). So far, several trust-based CF approaches have been proposed to overcome data sparsity and cold-start problems as well as to increase recommendable items (Ghavipour & Meybodi, 2016; Moradi, Ahmadian, & Akhlaghian, 2015; P. Massa & Avesani, 2007; Ranjbar Kermany & Alizadeh, 2017). Trust statements can be explicitly collected from users or can be implicitly inferred from users behaviors (S. Ahmadian, M. Meghdadi, & Afsharchi, 2018a; S. Ahmadian, P. Moradi, & Akhlaghian, 2014). Liu and Lee proposed a specific approach which does not directly use the trust information; instead they take into account the number of exchanged messages among the users of the system to construct the trust network (Liu & Lee, 2010). Alahmadi and Zeng presented a framework to apply short texts posted by users friends in microblogs as an additional data source to build the trust network (Alahmadi & Zeng, 2015). Since explicit trust statements are directly specified by the users, they are more accurate and reliable than implicit ones in determining social relationships among users (Cho, Kwon, & Park, 2009; Ingoo, Kyong, & Tae, 2003; Lathia, Hailes, & Capra, 2008; Manolopoulus, Nanopoulus, Papadopoulus, & Symeonidis, 2008).

The research In (Abdul-Rahman & Hailes, 2000) has been shown that a user constructs his/her social connections with someone who has similar tastes. Massa and Avesani showed that adding social network data to traditional collaborative filtering improves the recommendation results (P. Massa & Avesani, 2007). Gharibshah and

Jalili studied the relation between RSs and connectedness of users-items bipartite interaction network (Gharibshah & Jalili, 2014). Guo et al. proposed a method which merged the ratings of users trusted neighbors with the other information sources to identify their preferences (G. Guo, J. Zhang, & Thalmann, 2014). Yang et al. proposed a Bayesian inference based recommendation method for online social networks (X. Yang, Y. Guo, & Liu, 2013). In this method, the similarity value between each pair of users is measured using a set of conditional probabilities derived from their mutual ratings. Jiang et al. introduced a framework to incorporate interpersonal influences of users in social network with their individual preferences to improve the accuracy of social recommendation (Jiang, Cui, Wang, Zhu, & Yang, 2014).

Purchase/rating time is one of the most important contextual information that can be used to design RSs with high precision (Xiong, Chen, Huang, Schneider, & Carbonell, 2010). The main motivation for time-aware RS is that in realistic scenarios users tastes might change over time.

3. Methodology

We propose a time and trust-aware RS using a graph-based community detection method consists of four steps: 1: developing a user-item rating matrix, 2: constructing a time weighted user-user network, 3: performing graph- based community detection, 4: recommending Top-N items. In the first step, the user-item rating matrix is developed by adding some implicit ratings and the quality of the implicit ratings is evaluated using a reliability measurement. In the second step, a time-weighted user-user network is constructed based on the combination of trust relationships and similarity between users. Moreover, the timestamps of user-item ratings are considered to calculate the similarity between users. In the third step, a graph-based community detection method classifies similar users into appropriate communities. Finally, in the fourth step, it predicts the rating for each unobserved item and top-N recommendations is generated for the target user.

We proposed a new community detection method that consists of three phases. First, the initial centers of communities are obtained using a sparsest subgraph of weighted user-user network. It should be

noted that the initial centers must have the maximum dissimilarities with each other based on the general concept of clustering and community detection algorithms. Then users can be assigned to their nearest communities. For each user proposed system calculated the fitness function. User has associated to community which has high value of fitness function. Then the centers of communities were updated in order to maximize a fitness function. This process is iteratively repeated until members of communities do not change and steady state is achieved. A set of communities are identified where the users are assigned to their corresponding communities. Some of the communities may have overlap and they can be merged. The final communities were used as the nearest neighbors set of the active user in the same community for the recommendation.

4. Conclusion

Our proposed algorithm solves the sparsity of rating matrix by adding the implicit rating and solved cold-start problem for new users by considering the trust between the users. We applied the proposed algorithm on extended Epinions dataset and compared its performance with similar algorithms. The experimental results showed that our proposed algorithm outperforms the other algorithms according to the accuracy and recommends the top@N items with high precision.

Keywords: Recommender Systems, Time, Trust, Community Detection, Collaborative Filtering.

مقاله پژوهشی تاریخ ارسال: ۱۴۰۱/۰۸/۰۳ تاریخ بازنگری: ۱۳۹۹/۰۶/۰۴ تاریخ پذیرش: ۱۴۰۲/۰۹/۱۲
ISSN: 2821-0964 eISSN: 2821-0816

مطالعات مدیریت کسب و کار هوشمند
سال دوازدهم، شماره ٤٦، زمستان ١٤٠٢، ص ٣٢٧ تا ٣٦٠
ims.atu.ac.ir
DOI: 10.22054/ims.2023.70722.2247

سیستم توصیه‌گر زمان و اعتماد مبنا با استفاده از تشخیص جوامع مبتنی بر گراف

فاطمه رضایی‌مهر دانشجوی دکتری رشته مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیکز، دانشگاه صنعتی خواجه‌نصیرالدین طوسی، دانشکده کامپیوتر، گروه هوش مصنوعی تهران، ایران.

چیترا دادخواه * استادیار دانشگاه صنعتی خواجه‌نصیرالدین طوسی، دانشکده کامپیوتر، گروه هوش مصنوعی، تهران، ایران

چکیده

اخیراً، فضای اینترنت نقش زیاد و بااهمیتی در زندگی انسان‌ها پیدا کرده است؛ اما محتوای موجود در محیط جهانی وب باید با توجه به نیازهای روز کاربران باشد و اطلاعات مفید و به‌روز و مناسب با سلیقه کاربر را در اختیار آن‌ها قرار دهد. در این راستا، سیستم‌های توصیه‌گر به کاربران کمک کرده و اقلامی را که به سلیقه کاربران نزدیک‌تر است را به آن‌ها در کمترین زمان ممکن توصیه می‌کند. امروزه، با افزایش مقدار داده، استفاده از سیستم‌های توصیه‌گر رو به افزایش است. از طرفی دیگر این سیستم‌ها با چالش‌هایی ازجمله تغییر سلیقه کاربران در طی زمان، شروع سرد، خلوت بودن ماتریس کاربر-قلم، حملات توسط افراد جعلی در سیستم و تأثیر منفی آن‌ها در لیست توصیه سیستم روبرو هستند. هدف این مقاله ارائه یک سیستم توصیه‌گر زمان و اعتماد مبنا جهت بهبود کارایی و افزایش دقت توصیه‌های سیستم است. سیستم پیشنهادی در ابتدا با افزودن برخی امتیازهای ضمنی قابل‌اعتماد به ماتریس امتیازدهی کاربر- قلم، مشکل پراکندگی داده را حل نموده و سپس یک شبکه وزن‌دار کاربر-کاربر بر اساس زمان ارائه نظر کاربر نسبت به قلم و روابط اعتماد میان کاربران تولید می‌نماید که بدین ترتیب مشکل شروع سرد و تغییر سلیقه کاربر در طی زمان را رفع می‌کند. سیستم توصیه‌گر پیشنهادی بر اساس الگوریتم تشخیص جامعه جدیدی که در این مقاله ارائه شده است، نزدیک‌ترین کاربران همسایه و هم‌سلیقه با کاربر فعال را پیدا نموده و بر اساس روش پالایش همکارانه، کا-بالاترین قلم را به کاربر پیشنهاد می‌دهد. نتایج ارزیابی سیستم پیشنهادی برای

* نویسنده مسئول : dadkhah@ kntu.ac.ir

سیستم توصیه‌گر مبتنی بر فیلم بر روی مجموعه داده Epinions نشان می‌دهد که سیستم پیشنهادی نسبت به سیستم‌های پایه از کارایی بالاتری برخوردار است.

کلیدواژه‌ها: سیستم‌های توصیه‌گر، زمان، اعتماد، تشخیص جوامع، پالایش همکارانه.

مقدمه

داده‌کاوی[1] یکی از موضوعات مهم در حوزه‌های مختلف تحقیقاتی ازجمله شبکه‌های اجتماعی و خریدهای برخط[2] تبدیل شده است. حجم زیادی از داده‌های دیجیتال اغلب توسط کاربران هنگام خرید/ارزیابی اقلام تولید می‌شود. چنین اطلاعاتی همراه با ماتریس کاربر-قلم[3] می‌تواند برای مدل‌سازی ترجیحات کاربر و ارائه توصیه‌های مناسب به آن‌ها استفاده شود. سیستم‌های توصیه‌گر[4] برای انجام چنین وظایفی توسعه یافته‌اند و یکی از موضوعات اصلی علوم کامپیوتر در سال‌های اخیر بوده که بر روی داده‌های موجود در شبکه‌های اجتماعی در مقیاس بزرگ کار می‌کنند. سیستم‌های توصیه‌گر در بسیاری از زمینه‌ها مانند فیلم، ویدیو، صفحات وب، برنامه‌های کاربردی، بسته‌های مسافرتی، گردشگری، تجارت الکترونیک، مدل‌های فرآیند کسب‌وکار، مشاغل و غیره در کاربردهای صنعتی و دانشگاهی به‌صورت موفقیت‌آمیزی در سال‌های اخیر مورداستفاده قرار گرفته‌اند.

ازآنجاکه توصیه‌ها معمولاً در سیستم‌های توصیه‌گر شخصی‌سازی می‌شوند، کاربران یا گروه‌های کاربری مختلف توصیه‌های مختلفی دریافت می‌کنند. برای توصیه‌های شخصی، سیستم فهرستی از اقلام رتبه‌بندی شده را ارائه می‌دهد. در انجام این رتبه‌بندی، سیستم تلاش می‌کند تا بر اساس اولویت‌ها و محدودیت‌ها، مناسب‌ترین و بهترین محصولات و خدمات را برای کاربران که منطبق با سلیقه آن‌ها باشد، پیش‌بینی کند و برای تکمیل چنین عملیاتی، تراکنش‌های گذشته کاربران را جمع‌آوری می‌کند (Ricci, 2010). سیستم‌های توصیه‌گر از منابع داده‌ای مختلف مانند ماتریس امتیازات، ویژگی‌های کاربر و اقلام برای استخراج همبستگی بین کاربران و اقلام استفاده می‌کنند. بر اساس تکنیک و روش‌های توصیه در (Aggarwal, 2016)، سیستم‌های توصیه‌گر را می‌توان به مبتنی بر

1. Data Mining
2. Online
3. Item
4. Recommender Systems

محتوا (CB)[1]، پالایش همکارانه (CF)[2]، مبتنی بر دانش (KB)[3]، مبتنی بر اطلاعات دموگرافیک (DB)[4] و سیستم ترکیبی (HR)[5] تقسیم کرد. مدل پالایش همکارانه توصیه‌هایی که به کاربران فعال ارائه می‌دهد بر اساس شباهت سنجی میان کاربران باتوجه به نظرات تمامی کاربران سیستم نسبت به اقلام مشترک امتیاز داده‌شده با نظر کاربر فعال، است. در این نوع سیستم‌ها تمرکز بر روی نظرات کلیه کاربران سیستم است. در رویکرد مبتنی بر محتوا شباهت سنجی بر اساس ویژگی‌های اقلام در کنار نظر خود کاربر فعال در نظر گرفته‌شده و نظرات سایر کاربران تأثیری در توصیه سیستم نخواهد داشت. در مدل مبتنی بر دانش کاربران به‌صورت صریح نیازهای خود را اعلام نموده و توصیه‌های سیستم منطبق با این نیازها ارائه می‌گردد و دیگر نیازی به نظرات سایر کاربران برای یافتن سلیقه کاربر فعال و ارائه توصیه‌ها بر آن اساس نیست. این رویکرد معمولاً برای اقلامی بکار گرفته می‌شود که اکثریت کاربران نسبت به آن‌ها تراکنشی انجام نمی‌دهند مثل خرید وسایل آنتیک یا خرید منزل. به‌عنوان‌مثال کاربر فعال برای خرید منزل به‌طور صریح سلیقه خود را در قالب ویژگی‌های منزل مانند یک خانه با دو اتاق‌خواب، با حداکثر بودجه، ۱ حمام و غیره به‌صورت صریح عنوان می‌کند. این در حالی است که در سیستم‌های مبتنی بر پالایش همکارانه فقط امتیاز نظر کاربر فعال در اقلام مشاهده‌شده وجود دارد و هیچ اطلاعات اضافی در رابطه با سلیقه او در سیستم وجود ندارد و نیاز است که با توجه به پیدا کردن کاربران هم سلیقه با او توصیه‌هایی ارائه گردد. در رویکرد مبتنی بر دموگرافیک پروفایل کاربران حاوی سن و جنسیت و شغل و غیره در توصیه‌ها دخالت داده می‌شوند. هرکدام از این روش‌ها دارای نقاط قوت و ضعف می‌باشند که با ترکیب این رویکردها در مدل ترکیبی از نقاط قوت آن‌ها در بهبود دقت و عملکرد سیستم توصیه‌گر استفاده می‌شود.

---

1. Content-Based
2. Collaborative Filtering
3. Knowledge-Based
4. Demographic-Based
5. Hybrid System

امروزه مدل توسعه‌یافته دیگری تحت عنوان مبتنی بر زمینه[1] که در کنار ماتریس امتیازدهی کاربر-قلم از اطلاعات اضافی مانند مکان جغرافیایی، زمان، نظرات متنی کاربران، ارتباطات اجتماعی و غیره در توصیه‌های سیستم استفاده می‌نماید، کاربرد وسیعی را پیدا کرده است (AGGARWAL, 2016).

اگرچه استفاده از سیستم‌های توصیه‌گر موفقیت‌آمیز بوده است، اما مشکلات متعددی مانند خلوت بودن ماتریس امتیازات، شروع سرد، مقیاس‌پذیری و حمله رنج می‌برند. وقتی کاربران به تعداد کمی از اقلام امتیاز می‌دهند، مشکل خلوت بودن ماتریس امتیازات رخ می‌دهد و درنتیجه، همسایگان کاربر فعال را نمی‌توان به‌درستی شناسایی کرد. مشکل شروع سرد برای کاربران جدید یا اقلامی رخ می‌دهد که به‌تازگی وارد سیستم شده‌اند و هیچ سابقه‌ای از امتیازات کاربر جدید و یا نظرات کاربران برای اقلام جدید وجود ندارد؛ بنابراین معیارهای شباهت سنجی یا غیرممکن است یا به نتایج غیرقابل‌اعتماد منجر می‌شوند. کامنت‌های متنی، روابط دوستی و اعتماد رویکردهایی برای حل مشکل شروع سرد در شبکه‌های اجتماعی هستند. همان‌طور که سیستم‌های توصیه‌گر مبتنی بر پالایش همکارانه به‌راحتی قابل‌درک و فهم هستند و کاربردهای بسیار زیادی دارند اما در مقابل حملات آسیب‌پذیر می‌باشند که اخیراً روش‌هایی برای شناسایی و حذف این حملات توسط محققین معرفی شده‌اند که در (Rezaimehr & Dadkhah, 2021) موردمطالعه قرار گرفته‌اند.

در این مقاله، از روابط اعتماد و زمان امتیازدهی جهت طراحی پیشنهادی سیستم با دقت بالا ارائه شده است. سیستم توصیه‌گر پیشنهادی ابتدا یک نمودار شباهت وزن‌دار کاربر-کاربر را ساخته و سپس خلوت‌ترین زیرگراف را شناسایی و ساختار جامعه کاربران را بر اساس یک الگوریتم جدید شناسایی جامعه ارائه‌شده در این مقاله به دست می‌آورد که برای آن از اطلاعات زمان و اعتماد میان کاربران استفاده نموده است. درنهایت بر اساس کا- نزدیک‌ترین همسایه[2] به کاربر فعال کا-بالاترین قلم[1] را پیشنهاد می‌دهد. سیستم

1. Context-based
2. K Nearest Neighborhood (KNN)

پیشنهادی از معیار قابلیت اطمینان برای کاهش حجم پراکندگی و خلوت بودن ماتریس امتیازات و همچنین از یک روش جدید تشخیص جامعه همپوشان در فرآیند توصیه خود استفاده می‌کند.

ساختار مقاله به‌صورت زیر سازمان‌دهی شده است: کارهای پیشین در زمینه سیستم‌های توصیه‌گر زمان و اعتماد مبنا در بخش ۲ بررسی می‌شوند. در بخش ۳ ساختار سیستم پیشنهادی ارائه می‌گردد. بخش ۴، نتایج ارزیابی سیستم پیشنهادی بر روی یک مجموعه داده واقعی بیان شده است. درنهایت، بخش ۵ نتیجه‌گیری و کارهای آینده را ذکر می‌کند.

پیشینه پژوهش

در سال‌های اخیر، محققان تلاش کرده‌اند تا دقت توصیه‌های سیستم توصیه‌گر طراحی‌شده را برای حفظ کاربران و افزایش سود، افزایش دهند. امروزه سیستم‌های توصیه‌گر از حالت سنتی خود خارج شده‌اند و محققین بر این باورند که فرمول‌بندی آن به‌عنوان یک مسئله تصمیم‌گیری می‌تواند سبب تعامل بهتر کاربر-سیستم شود (Afsar, Crump, & Far, 2022). در بسیاری از روش‌های اخیر ارائه‌شده، تکنیک‌های شبکه عصبی گراف(GNN[۲]) به‌طور گسترده مورداستفاده قرار گرفته است. دلیل استفاده گسترده این است که محققین بر این باور بوده‌اند که بیشتر اطلاعات در سیستم‌های توصیه‌گر اساساً ساختار نموداری دارند و GNNدر یادگیری نمایش نمودار برتری دارد (Wu, Sun, Zhang, Xie, & Cui, 2022). برخی از محققین بر روی بهینه‌سازی عملکرد سیستم‌های پیشنهادی خود با استفاده از الگوریتم‌های تکاملی کار کرده‌اند (Tohidi & Dadkhah, 2020) و برخی دیگر از اطلاعات اضافی مانند زمان، مکان و غیره استفاده نموده‌اند. در این بخش مروری بر کارهای مشابه با سیستم پیشنهادی ارائه می‌نماییم.

---

1. Top-K Items
2. Graph Neural Network

سیستم‌های توصیه‌گر مبتنی بر اعتماد اخیراً به جامعه علوم رایانه معرفی شده‌اند. در این نوع سیستم‌ها، امتیازات کاربران مورد اعتماد با کاربر هدف در شبکه‌های اجتماعی نیز در فرآیند توصیه در نظر گرفته می‌شود. مطالعات اخیر نشان داده است که ترکیب عوامل اجتماعی مانند اعتماد میان کاربران منجر به بهبود کیفیت توصیه در سیستم‌های توصیه‌گر می‌گردد.

تاکنون چندین رویکرد مبتنی بر پالایش همکارانه مبتنی بر اعتماد برای غلبه بر پراکندگی داده‌ها و مشکلات شروع سرد برای افزایش قابلیت اطمینان لیست توصیه‌ها پیشنهاد شده است (Ghavipour & Meybodi, 2016; Massa & Avesani, 2007; Moradi et al., 2015; Ranjbar Kermany & Alizadeh, 2017).

اعتماد را می‌توان به‌صراحت از کاربران جمع‌آوری کرد یا به‌طور ضمنی از رفتارهای کاربران مانند اطلاعات امتیازدهی استنباط نمود. احمدیان و همکارانش در روش ارائه داده‌شده، یک شبکه اعتماد برای هر کاربر بر اساس امتیازات او در سیستم ایجاد می‌نمایند (Ahmadian et al., 2020). لئو و لی رویکرد خاصی را پیشنهاد کردند که مستقیماً از اطلاعات اعتماد استفاده نمی‌کند. در عوض آن‌ها تعداد پیام‌های مبادله شده بین کاربران سیستم را در نظر می‌گیرند (Liu & Lee, 2010). الامردی و همکارش چارچوبی را برای اعمال متون کوتاه ارسال‌شده توسط دوستان کاربران در میکروبلاگ‌ها را به‌عنوان منبع داده اضافی برای ایجاد شبکه اعتماد معرفی کردند (Alahmadi & Zeng, 2015). هر چه امتیازاتی که دو کاربر به یک قلم مشترک داده‌اند در شبکه به هم نزدیک‌تر باشند، ارزش آن بیشتر و بنابراین اعتماد میان آن‌ها (درواقع سلیقه آن‌ها) بیشتر خواهد بود. از سوی دیگر، در روش‌های صریح اعتماد از ارتباطات اجتماعی از پیش تعیین‌شده در میان کاربران یک شبکه اجتماعی استفاده می‌کنند. ازآنجایی‌که اعتماد صریح مستقیماً توسط کاربران مشخص می‌شود، در تعیین روابط اجتماعی دقیق‌تر و قابل‌اعتمادتر از ضمنی هستند.

کارهای اخیر پیشنهادشده با در نظر گرفتن کاربران شبکه‌های اجتماعی برخط در سیستم‌های توصیه‌گر آگاه از اعتماد ارائه شده است. این شبکه‌ها به کاربران اجازه می‌دهند

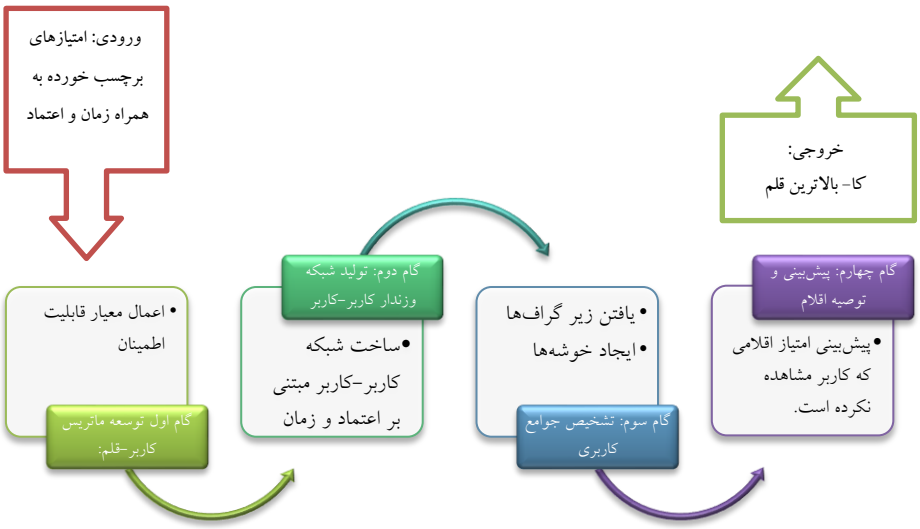
تا روابط اعتماد صریح مشابه آنچه در دنیای واقعی وجود دارد ایجاد کنند. در (Abdul-Rahman & Hailes, 2000) نشان داده شده است که کاربران با سلیقه‌های مشابه با یکدیگر ارتباط برقرار می‌کنند. ماسا و آوسانی نشان دادند که افزودن داده‌های شبکه اجتماعی به رویکرد پالایش همکارانه سنتی نتایج توصیه‌ها را بهبود می‌بخشد(Massa & Avesani, 2007). گوو و همکارانش روشی را پیشنهاد کردند که امتیازات همسایگان مورد اعتماد کاربران را با سایر منابع اطلاعاتی برای شناسایی سلایق آن‌ها ادغام کردند(Guo, Zhang, & Thalmann, 2014). مرادی و احمدیان چارچوبی را برای ترکیب اعتماد ضمنی و صریح پیشنهاد کردند (Moradi & Ahmadian, 2015). آن‌ها از معیارهای قابلیت اطمینان برای ارزیابی چگونگی برآورد اعتماد ضمنی استفاده کردند. یانگ و همکارانش یک روش توصیه مبتنی بر استنتاج بیزین برای شبکه‌های اجتماعی برخط ارائه نمودند (Yang, Guo, & Liu, 2012).

انگیزه اصلی سیستم توصیه‌گر آگاه از زمان این است که در سناریوهای واقعی، سلیقه کاربران ممکن است در طول زمان تغییر کند. اکثر الگوریتم‌های توصیه‌گر سنتی از داده‌ها به‌صورت نظرات ثابت کاربران بهره‌برداری می‌کنند. اخیراً در رویکردهای مبتنی بر پالایش همکارانه تغییر سلیقه کاربران در فرآیندهای توصیه توسط الگوریتم‌های مختلفی در نظر گرفته شده است. در الگوریتم‌های پنجره زمانی، تنها تعدادی از جدیدترین امتیازها در فرآیند توصیه در نظر گرفته می‌شوند. در روش‌های مبتنی بر زمان، اثر زمانی داده‌ها با افزایش امتیازهای اخیر و کم کردن وزن امتیازهای‌های قدیمی‌تر کنترل می‌شود. زنگ و همکارش از یک تابع برای محاسبه اهمیت و سودمندی برچسب زمان تعامل استفاده کردند (Zheng & Li, 2011). کرن یک روش فاکتورسازی ماتریسی که شامل اثرات زمانی است برای سیستم توصیه‌گر معرفی نموده است (Koren, 2010). دانشمند و همکارانش نیز یک سیستم توصیه‌گر آگاه از زمان که وابستگی‌های بین اقلام را با استفاده از انتشار آماری مدل‌سازی کرده، ارائه دادند(Daneshmand, Javari, Abtahi, & Jalili, 2015). اخیراً با ظهور محاسبات شخصیت به‌عنوان یک زمینه تحقیقاتی جدید مرتبط با هوش مصنوعی و

روانشناسی شخصیت، شاهد گسترش بی‌سابقه سیستم‌های توصیه‌گر مبتنی بر شخصیت بوده‌ایم. برخلاف سیستم‌های توصیه گر مرسوم، این سیستم‌های جدید مشکلات سنتی مانند شروع سرد و مشکلات پراکندگی داده را حل می‌کنند (Hui, Zhang, Zhou, Wen, & Nian, 2022). بنسال و همکارانش در (Bansal, Flannery, & Woolhouse, 2021) با بررسی پراکندگی سبک موردعلاقه کاربران در میان مجموعه‌های موسیقی تلفن همراه افراد، رفتار برگرفته از بارگزاری[1] موسیقی، زیرگروه‌های کاربری مبتنی بر سبک‌های مختلف را تجزیه‌وتحلیل کردند.

روش پیشنهادی

همان‌طور که در شکل ۱ مشاهده می‌نمایید، سیستم پیشنهادی از چهار گام اصلی تشکیل شده است: ۱: توسعه ماتریس کاربر- قلم، ۲: ساخت شبکه وزن‌دار کاربر –کاربر، ۳: شناسایی جوامع، ۴: پیش‌بینی و توصیه کا-بالاترین قلم[2].



شکل ۱. ساختار سیستم پیشنهادی

1. Download
2. Top@k items

گام اول: توسعه ماتریس کاربر-قلم

در بسیاری از سیستم‌های واقعی، ماتریس امتیازدهی کاربر- قلم بسیار پراکنده است، زیرا کاربران عمدتاً تمایلی به ثبت امتیاز برای اقلام ندارند. هدف از این مرحله توسعه ماتریس امتیازدهی کاربر- قلم به‌منظور غلبه بر این مشکل است. برای این منظور در ابتدا با استفاده از رابطه ۱، امتیاز اقلام مشاهده نشده توسط کاربران پیش‌بینی می‌شود.

$$IR_{a,i} = \bar{r}_a + \frac{\sum_{u \in N_{a,i}} sim_{a,u}(r_{u,i} - \bar{r}_u)}{\sum_{u \in N_{a,i}} sim_{a,u}} \quad (۱)$$

که $\bar{r}_a$ میانگین امتیازات کاربر a، $N_{a,i}$ مجموعه‌ای از کاربران همسایه‌ی موجود در شبکه‌ی اعتماد کاربر a که به قلم i امتیاز داده‌اند، می‌باشد. $sim_{a,u}$ تشابه میان کاربر u و a با استفاده از ضریب همبستگی پیرسون طبق رابطه ۲ محاسبه می‌شود.

$$sim_{a,u} = \frac{\sum_{i \in I_{a,u}} \left(r_{a,i} - \bar{r}_a\right)\left(r_{u,i} - \bar{r}_u\right)}{\sqrt{\sum_{i \in I_{a,u}} \left(r_{a,i} - \bar{r}_a\right)^2} \sqrt{\sum_{i \in I_{a,u}} \left(r_{u,i} - \bar{r}_u\right)^2}} \quad (۲)$$

$I_{a,u}$ شامل کلیه اقلامی هست که توسط دو کاربر a,u امتیازدهی شده‌اند.

برای محاسبه این مقادیر، از کل امتیازات موجود در ماتریس امتیازدهی کاربر-قلم استفاده شده است. ممکن است امتیازات دو کاربری که اصلاً هم سلیقه نباشد، بر این مقدار تأثیر بگذارد. این امر فقط جهت پر کردن ماتریس کاربر-قلم انجام شده است و ممکن است برخی از امتیازات پیش‌بینی‌شده واقعاً قابل‌اتکا نباشد. درنتیجه باید این مقادیر درستی آزمایی شوند. برای این مهم از یک معیار اطمینان (Hernando, Bobadilla, Ortega, & Tejedor, 2013) شامل فاکتور مثبت (fp) و منفی (f_n) برای محاسبه میزان اعتماد امتیازهای پیش‌بینی‌شده استفاده می‌نماید. این مقدار با استفاده از رابطه ۳ محاسبه می‌شود.

$$R_{a,i} = \left(fp(r_{a,i}).\, f_n(r_{a,i})^{fp(r_{a,i})}\right)^{\frac{1}{1+fn(r_{a,i})}} \quad (۳)$$

فاکتور مثبت تأثیر مثبت امتیازات خود کاربر فعال و همسایگان او را محاسبه می‌کند. درواقع بر اساس جمع شباهت بین کاربر هدف و نزدیک‌ترین همسایگان آن با استفاده از

رابطه ۴ محاسبه می‌شود.

$$f_s\left(S_{a,i}\right)=1-\frac{\bar{S}}{\bar{S}+S_{a,i}} \quad (۴)$$

$\bar{S}$ مقدار میانه‌ی مقادیر $S_{a,i}$ برای سیستم توصیه‌گر می‌باشد که از رابطه ۵ محاسبه می‌گردد.

$$S_{a,i}=\sum_{u\in N_{a,i}} sim_{a,u} \quad (۵)$$

فاکتور منفی، تأثیر منفی امتیازات کاربر فعال را محاسبه می‌کند و بر اساس واریانس امتیازهای همسایگان کاربر فعال با استفاده از روابط ۶، ۷ و ۸ محاسبه می‌شود.

$$f_V(V_{a,i})=\left(\frac{max-min-V_{a,i}}{max-min}\right)^{\gamma} \quad (۶)$$

$$V_{a,i}=\frac{\sum_{u\in K_{a,i}} sim_{a,u}.(r_{u,i}-\bar{r_u}-IR_{a,i}+\bar{r_a})^2}{\sum_{u\in K_{a,i}} sim_{a,u}} \quad (۷)$$

$$\gamma=\frac{\ln 0.5}{\ln\frac{max-min-\bar{\bar{V}}}{max-min}} \quad (۸)$$

$\bar{\bar{V}}$مقدار میانه‌ی مقادیر $V_{a,i}$، max و min به ترتیب ماکزیمم و مینیمم امتیازات داده‌شده به اقلام سیستم می‌باشند.

درنهایت اگر مقدار قابلیت اطمینان محاسبه‌شده در رابطه ۳ از یک حد آستانه φ بزرگ‌تر باشد، سیستم پیشنهادی آن‌ها به ماتریس کاربر- قلم اضافه می‌نماید.

گام دوم: تولید شبکه وزن‌دار کاربر-کاربر

اگر دو کاربر $\mathbf{u}$ و $\boldsymbol{v}$ به قلم‌های مشترکی علاقه نشان داده باشند، بین آن دو کاربر در ابتدا یک یال ایجاد نموده و میزان تشابه بین دو کاربر $\mathbf{u}$ و $\boldsymbol{v}$ با استفاده از ضریب همبستگی پیرسون و با در نظرگرفتن زمان امتیازدهی محاسبه و به‌عنوان وزن یال در نظر گرفته می‌شود. استفاده از زمان امتیازدهی، موجب می‌شود که تغییر سلیقه کاربر در طی زمان نیز

لحاظ شود.

با این فرض که کاربر a به قلم i در دوره زمانی t امتیاز داده است، بر طبق رابطه ۹ وزن زمانی این امتیازدهی برای زمان T که نشانگر زمان ارائه توصیه به کاربر در نظر گرفته شده است، محاسبه می‌گردد. تمامی مقادیر محاسبه‌شده در یک ماتریس $W = [w_{ij}]$ ذخیره می‌شوند.

$$W_{a,i}^{t} = \begin{cases} e^{-(T-t)/\theta} & if\ t > 0 \\ 0 & otherwise \end{cases} \quad (۹)$$

$\boldsymbol{\theta}$ یک پارامتر در بازه $\left[\mathbf{0.25}\ ,\ \mathbf{1.25}\right]$ است که توسط کاربر تعیین می‌شود. باید توجه داشت هرچه مقدار این پارامتر افزایش یابد، تأثیر زمان کمتر می‌شود. بر همین اساس ماتریس امتیازات جدید $\mathbf{tr}$ با استفاده از ترکیب ماتریس کاربر–قلم اولیه و ماتریس $\mathbf{W}$ با استفاده از رابطه ۱۰ ساخته می‌شود.

$$\mathrm{tr}_{\mathrm{a,i}} = \begin{cases} \mathrm{r}_{\mathrm{a,i}} \times \mathrm{W}_{\mathrm{a,i}}^{\mathrm{t}}\ \mathrm{if} \mathrm{r}_{\mathrm{a,i}} > 0 \\ 0\ \mathrm{otherwise} \end{cases} \quad (۱۰)$$

$\mathbf{r_{a,i}}$ امتیازی است که کاربر $\mathbf{a}$ به قلم $\mathbf{i}$ داده و در ماتریس کاربر–قلم اولیه وجود دارد. حالا سیستم پیشنهادی با استفاده از مقادیر ماتریس جدید، تشابه میان کاربران از بعد زمان امتیازدهی را با استفاده از رابطه ۱۱ محاسبه می‌نماید.

$$\mathrm{t_sim}_{\mathrm{a,u}} = \frac{\sum_{\mathrm{i}\in \mathrm{I}_{\mathrm{a,u}}} (\mathrm{tr}_{\mathrm{a,i}} - \overline{\mathrm{tr}}_{\mathrm{a}})(\mathrm{tr}_{\mathrm{u,i}} - \overline{\mathrm{tr}}_{\mathrm{u}})}{\sqrt{\sum_{\mathrm{i}\in \mathrm{I}_{\mathrm{a,u}}} (\mathrm{tr}_{\mathrm{a,i}} - \overline{\mathrm{tr}}_{\mathrm{a}})^2} \sqrt{\sum_{\mathrm{i}\in \mathrm{I}_{\mathrm{a,u}}} (\mathrm{tr}_{\mathrm{u,i}} - \overline{\mathrm{tr}}_{\mathrm{u}})^2}} \quad (۱۱)$$

که می‌توان گفت $\overline{\mathbf{tr}}_{\mathbf{a}}$ میانگین امتیازات زمانی کاربر $\mathbf{a}$ است و $\mathbf{I_{a,u}}$ اقلام مشابهی است که توسط هردو کاربر u و a امتیازدهی شده است.

درنهایت تشابه میان کاربران از ترکیب رابطه اعتماد و تشابه زمانی، با استفاده از رابطه ۱۲ به دست می‌آید.

$$H_{a,u}=\begin{cases}\frac{2\times t_sim_{a,u}\times T_{a,u}}{t_sim_{a,u}+T_{a,u}} t_sim_{a,u}>0 \text{ and} T_{a,u}>0\\ t_sim_{a,u} t_sim_{a,u}>0 \text{ and} T_{a,u}\le 0\\ T_{a,u} t_{sim_{a,u}}\le 0 \text{ and} T_{a,u}>0\\ 0 t_{sim_{a,u}}\le 0 \text{ and} T_{a,u}\le 0\end{cases} \quad (۱۲)$$

$\mathbf{T_{a,u}}$اعتماد ضمنی میان کاربران u و a با استفاده از رابطه ۱۳ محاسبه می‌شود.

$$T_{a,u}=\frac{d_{max}-d_{a,u}+1}{d_{max}} \quad (۱۳)$$

سیستم پیشنهادی زمانی که اعتماد صریح بین دو کاربر a و u در مجموعه وجود نداشته باشد با توجه به رابطه ۱۳ اعتماد ضمنی بین آنها را محاسبه می‌نماید.

$\mathbf{d_{a,u}}$ فاصله میان کاربران u و a در شبکه اعتماد و $\mathbf{d_{max}}$ با استفاده از رابطه ۱۴ محاسبه می‌شود.

$$d_{max}=\frac{\ln(n)}{\ln(k)} \quad (۱۴)$$

که n و k به ترتیب اندازه و میانگین درجه کاربران در شبکه اعتماد است. با استفاده از **H** به تمام یال‌های موجود در شبکه کاربر-کاربر یک وزن اختصاص داده می‌شود.

گام سوم: تشخیص جوامع کاربری

در روش‌هایی که برای تشخیص جوامع استفاده می‌شوند، مهم‌ترین مشکل تشخیص اولیه تعداد خوشه‌ها و همچنین انتخاب تصادفی مراکز خوشه‌ها است. در این مرحله یک روش جدید تشخیص جوامع ارائه‌شده که گره‌هایی با دامنه ارتباط گسترده را به‌عنوان مرکز خوشه در نظر می‌گیرد. به عبارتی مرکز خوشه‌ها گره‌هایی با بیشترین لینک با سایر گره‌های موجود در شبکه تعیین می‌شوند. گراف $\mathbf{G}=(\mathbf{V},\mathbf{E},\mathbf{H})$ به‌عنوان شبکه وزن‌دار کاربر-کاربر است که V,E,H به ترتیب نمایش‌دهنده کاربران سیستم، روابط میان کاربران و وزن این روابط (طبق رابطه ۱۲) می‌باشد. چگالی زیر گراف $S\subseteq \mathbf{V}$ با استفاده از رابطه ۱۵ محاسبه می‌شود.

$$\rho(S)=\frac{\sum_{e\epsilon E(S)} H_e}{|S|} \quad ۱۵$$

که $\mathbf{H_e} \in \mathbf{H}$ وزن یال $\mathbf{e}$ و $\mathbf{E(S)}$ مجموعه یال‌های زیر گراف $\mathbf{S}$ را بیان می‌کند. وزن هر گره $\mathbf{i} \in \mathbf{S}$ با استفاده از رابطه ۱۶ محاسبه می‌شود.

$$\mathrm{wd_s(i)} = \sum_{\mathrm{e_{ij}} \in \mathrm{E(S)}} \mathrm{H_{e_{ij}}} \qquad ۱۶$$

$\mathbf{\widetilde{A}(S)}$ مجموعه گره‌های کاندید شامل گره‌هایی با درجه‌ای بالاتر از حد آستانه محاسبه‌شده توسط رابطه ۱۷ است.

$$((2 + 2\varepsilon) * \rho(\mathrm{S})) \qquad ۱۷$$

الگوریتم ۱، الگوریتم یافتن مرکز خوشه را بیان می‌نماید. الگوریتم، کاربران را بر اساس درجه مرتب می‌کند و کاربرانی که بالاترین درجه را دارند از مجموعه حذف می‌کند. این الگوریتم مرتباً تکرار می‌شود تا جایی که تعداد گره‌هایی که انتخاب می‌شوند از $\mathbf{k}$ کمتر باشد. درنتیجه خروجی الگوریتم، یک زیر گراف k گره‌ای از گراف اصلی است که در ابتدا هر گره به‌عنوان مرکز یک خوشه محسوب می‌شود. پارامتر K به‌عنوان تعداد خوشه‌ها در نظر گرفته می‌شود، ه باید مقدار آن به‌درستی تنظیم شود به‌گونه‌ای که نه آن‌قدر بزرگ باشد که تعداد زیادی خوشه با تعداد کمی گره در هر خوشه و نه آن‌قدر کم که تعداد کمی خوشه با تعداد بسیار گره در هر خوشه داشته باشیم.

الگوریتم ۱. یافتن مراکز خوشه‌ها

ورودی: $G = (V, E, H), K > 0$ و $\varepsilon > 0$
خروجی: $\{Init_{\tilde{C}}\}$ مراکز اولیه جوامع

۱. $\{\boldsymbol{Init_C}\} \leftarrow \boldsymbol{V}, \{\boldsymbol{Init_{\tilde{C}}}\} \leftarrow \emptyset$.
۲. تا زمانی که $\{\boldsymbol{Init_C}\} \neq \emptyset$ انجام بده
۳. $\{\widetilde{\boldsymbol{A}}(\boldsymbol{Init_C})\} \leftarrow \{\forall \boldsymbol{i} \epsilon \{\boldsymbol{Init_C}\} | \boldsymbol{wd_{Init_C}(i)} \geq \boldsymbol{(2 + 2\varepsilon) * \rho(\{Init_C\})}\}$;
۴. تمام $\boldsymbol{i} \in \{\widetilde{\boldsymbol{A}}(\boldsymbol{Init_C})\}$ بر اساس ترتیب نزولی $\boldsymbol{wd_{Init_C}(i)}$ مرتب کن.
۵. $\boldsymbol{r} = \frac{\boldsymbol{\varepsilon}}{\boldsymbol{1+\varepsilon}} \times |\{\widetilde{\boldsymbol{A}}(\boldsymbol{Init_C})\}|$;
۶. به تعداد r تا بالاترین گره‌های موجود در $\{\widetilde{\boldsymbol{A}}(\boldsymbol{Init_C})\}$ را به‌عنوان $\{\boldsymbol{A}(\boldsymbol{Init_C})\}$ انتخاب کنید.
۷. $\{\boldsymbol{Init_C}\} \leftarrow \{\boldsymbol{Init_C}\} - \{\boldsymbol{A}(\boldsymbol{Init_C})\}$
۸. اگر $|\{\boldsymbol{A}(\boldsymbol{Init_C})\}| == \boldsymbol{k}$ سپس

الگوریتم ۱. یافتن مراکز خوشه‌ها

۹. $\{\boldsymbol{Init_{\tilde{C}}}\} \leftarrow \{\boldsymbol{A(Init_C)}\}$
۱۰. خارج شو.
۱۱. در غیر این صورت اگر $|\{\boldsymbol{A(Init_C)}\}| > \boldsymbol{k\ and\ \rho}(\{\forall \boldsymbol{i} \epsilon \{\boldsymbol{A(Init_C)}\} | \boldsymbol{\rho(i)}\}) \leq$
سپس $\boldsymbol{AVG_\rho}(\{\boldsymbol{Init_{\tilde{C}}}\})$
۱۲. $\{\boldsymbol{Init_{\tilde{C}}}\} \leftarrow \boldsymbol{i}$
۱۳. اگر $|\{\boldsymbol{A(Init_{\tilde{C}})}\}| == \boldsymbol{k}$ سپس خارج شو.
۱۴. در غیر این صورت
۱۵. $\{\boldsymbol{Init_{\tilde{C}}}\} \leftarrow \{\boldsymbol{A(Init_C)}\}$
۱۶. پایان گام ۱۳
۱۷. پایان گام ۲
۱۸. return $\{\boldsymbol{Init_{\tilde{C}}}\}$

بعد از انتخاب مراکز اولیه جوامع، کاربران به نزدیک‌ترین خوشه اختصاص داده می‌شوند. روش پیشنهادی ارائه‌شده برای هر کاربر تابع برازش[1] را بر طبق رابطه ۱۷ به ازای تمام خوشه‌های موجود محاسبه کرده و کاربر را به خوشه‌ای که بیشترین تابع برازش را داشته باشد، اختصاص می‌دهد. در هر مرحله مرکز خوشه به‌روزرسانی می‌شود و تابعی با بالاترین مقدار برازش به‌عنوان مرکز خوشه انتخاب می‌شود.

$$f_{v_i} = \frac{\sum_{v_t \in C_j, v_t \neq v_i} H_{v_i,v_t}}{\sum_{v_t \in C_j, v_t \neq v_i} H_{v_i,v_t} + \sum_{C_k \in C, C_k \neq C_j} \sum_{v_t \in C_k, v_t \neq v_i} H_{v_i,v_t}} \quad ۱۷$$

در فرمول ۱۷ $\boldsymbol{C}$ مجموعه جوامع موجود است، $\boldsymbol{H_{v_i,v_t}}$، وزن بین دو گروه $\boldsymbol{v_t}$ و $\boldsymbol{v_i}$ است که با استفاده از فرمول ۱۲ شده است. این فرآیند تا جایی که اعضای جوامع تغییری نکنند، ادامه می‌یابد.

الگوریتم ۲ روش تشخیص جوامع معرفی‌شده در سیستم پیشنهادی را نشان می‌دهد. در ابتدا الگوریتم ۱ جهت تعیین مراکز خوشه‌ها فراخوانی می‌شود. جوامعی که باهم

1. Fitness Function

هم‌پوشانی دارند، می‌توانند باهم ادغام شوند. الگوریتم ۳ بیانگر الگوریتم ادغام جوامع است. در پایان این مرحله کاربران خوشه‌بندی شده‌اند و هر کاربر با همسایگان هم سلیقه خود در یک جامعه قرار داده می‌شوند.

الگوریتم ۲. تشخیص جوامع مبتنی بر گراف

ورودی: $G = (V, E, H), iter$.
خروجی: لیست جوامع $\{C\}$.

۱. با استفاده از الگوریتم ۱ مراکز اولیه جوامع را پیدا کن. $\{\boldsymbol{Init_{\tilde{C}}}\}$.
۲. $\boldsymbol{k' = |\{Init_{\tilde{C}}\}|}$.
۳. مجموعه $\boldsymbol{C_j = Init_{\tilde{C}_j}}$ به ازای مراکز خوشه $\forall \boldsymbol{j = 1, \dots, k'}$ که $\boldsymbol{C_j}$ همان جامعه $\boldsymbol{j}$ام می‌باشد.
۴. $\boldsymbol{v \leftarrow v - \{Init_{\tilde{C}}\}}$
۵. به ازای هر $\boldsymbol{iter}$ انجام بده
۶. برای هر $\boldsymbol{N \in v}$ تابع $\boldsymbol{f_N}$ را با استفاده از رابطه 17 برای $\boldsymbol{C_j}, \forall \boldsymbol{j = 1, \dots, k'}$ محاسبه کن.
۷. $j =$ *Community with* $\mathbf{argmax}\, \boldsymbol{f_N}$.
۸. $\boldsymbol{N}$ را به $\boldsymbol{C_j}$ اضافه کن.
۹. مرکز جامعه $\boldsymbol{C_j}$ را به‌روزرسانی کن.
۱۰. پایان گام ۶
۱۱. پایان گام ۵
۱۲. با استفاده از الگوریتم ۳ ادغام جوامع را محاسبه کن.
۱۳. $\{\boldsymbol{C}\}$ را برگردان.

الگوریتم ۳. ادغام جوامع هم‌پوشان.

ورودی: لیستی از جوامع $\{C\}$، ρ_1 و ρ_2.
خروجی: لیست نهایی جوامع $\{C'\}$.

۱. Calculate $\boldsymbol{\zeta = \frac{\rho_2 + \rho_2 \times \rho_1 - \rho_1}{\rho_1}}$
۲. برای $\boldsymbol{p, q \in C}$ و $\boldsymbol{p \neq q}$، $\boldsymbol{r_{pq} = \frac{\max(|C_p|, |C_q|)}{\min(|C_p|, |C_q|)}}$ را محاسبه کن.
۳. اگر $\boldsymbol{r_{pq} < \zeta}$ درنتیجه سایز دو جامعه $\boldsymbol{q}$ و $\boldsymbol{p}$ یکسان است و $\boldsymbol{\delta_{pq} = \beta \times \frac{|C_p \cap C_q|}{|C_p \cup C_q|} + (1 -}$

الگوریتم ۳. ادغام جوامع هم‌پوشان.

$\times \left(\frac{|NC_p \cap NC_q|}{|NC_p \cup NC_q|}\right)^{\beta}$ را محاسبه کن.

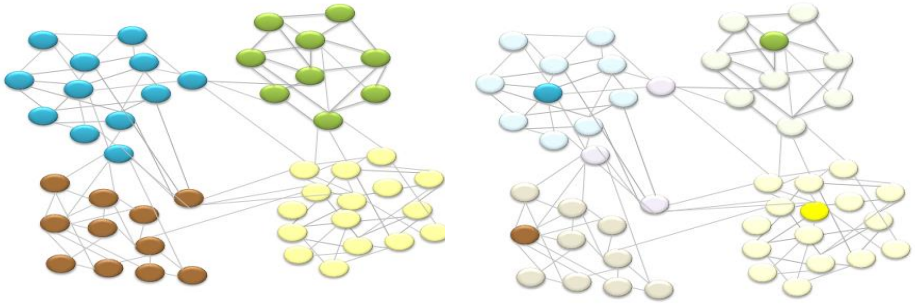
۴. اگر $r_{pq} \geq \zeta$ درنتیجه سایز دو جامعه q و p یکسان نیست و Calculate $\delta_{pq} = \frac{|C_p \cap C_q|}{\min(|C_p \cup C_q|)}$ را محاسبه کن.

۵. اگر $\delta_{pq} > \rho_2$ دو جامعه q و p ترکیب کن.

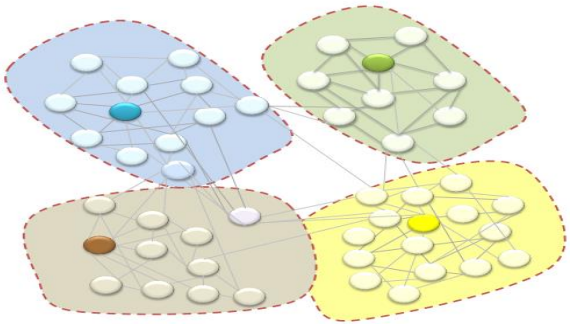
۶. $\{C'\}$ را برگردان.

در الگوریتم ۳، C_p و C_q جوامع هم‌پوشان pام و qام می‌باشند که $p \in \{1,..,K\}$ و $q \in$ $\{1,\ldots,K\}$. NC_p و NC_q مجموعه گره‌های همسایه C_p و C_q هستند که به‌صورت مستقیم با گره‌های جامعه در تماس هستند. $\rho_1, \rho_2 \in [0,1]$ پارامترهایی ورودی هستند. در شکل ۲ نمونه‌ای از عملکرد الگوریتم تشخیص جامعه را مشاهده می‌نمایید.

شکل ۲. نمونه‌ای از عملکرد الگوریتم تشخیص جوامع



(a) (b)



(c)

(a) شبکه اعتماد کاربر-کاربر، (b) مراکز جوامع، (c) جوامع نهایی.

گام چهارم: پیش‌بینی و توصیه اقلام

در این مرحله اقلامی که برای کاربر جذاب هستند، بر اساس خوشه‌بندی که در مرحله قبل به‌دست‌آمده است، بر اساس رابطه ۱۸ محاسبه می‌شود.

$$P_{a,i} = \overline{r_a} + \frac{\sum_{u \in N_{a,i}} H_{u,v} \left(r_{u,i} - \overline{r_u} \right)}{\sum_{u \in N_{a,i}} H_{u,v}} \qquad ۱۸$$

که در آن $\overline{r_a}$ میانگین امتیازات کاربر a و $N_{a,i}$ مجموعه‌ای از همسایگان کاربر a (کاربران موجود در جامعه‌ای که کاربر a به آن تعلق دارد) که به قلم i امتیاز داده‌اند و $H_{u,v}$ تشابه میان کاربر u و v که با استفاده از رابطه ۱۲ محاسبه‌شده است. پس از رتبه‌بندی مقادیر حاصله کا-بالاترین قلم به کاربر هدف توصیه داده می‌شود.

ارزیابی سیستم پیشنهادی

سیستم پیشنهادی به‌عنوان یک کاربرد سیستم توصیه‌گر مبتنی بر فیلم را انتخاب نموده است و بر این اساس بر روی مجموعه داده[1] Epinions که شامل اطلاعات اعتماد میان کاربران و زمان امتیازدهی فیلم مشاهده‌شده توسط ۲۲۱۶۶ کاربر، ۲۹۶۲۷۷ فیلم و ۹۲۲۲۶۸ امتیاز بین ۱ الی ۵ است، مورد آزمایش قرار گرفته است. امتیازات کاربران به مقاطع مختلف زمانی تقسیم‌بندی شده‌اند. همچنین نتایج روش پیشنهادی را با روش‌های پایه و مرسوم در سیستم‌های توصیه‌گر نظیر KMCF-U[2], TRACCF (Birtolo & Ronca, 2013), KMCF-I[3], TOTAR(Feng et al., 2015) مقایسه کرده‌ایم.

شش معیار ارزیابی Precision، Recall، F-measure، MAE[4]، RC[5] و RMSE[6] طبق

1. Epinions and Ciao Datasets (msu.edu)
2. User-Based k-means Collaborative Filtering
3. Item-Based k-means Collaborative Filtering
4. Mean Absolute Error
5. Rate Coverage
6. Root Mean Square Error

روابط ۱۹ تا ۲۴ موردبررسی قرار گرفته است. آزمایش‌ها بر روی یک ماشین با CPU core-i7 با 8GB RAM توسط زبان برنامه‌نویسی C# پیاده‌سازی شده است.

$$Precision = \frac{1}{n}\sum_{i=1}^{n}\frac{TP}{TP+FP} \quad ۱۹$$

(TP): مثبت درست، اقلامی که برای کاربران جذاب است و به آن‌ها توصیه می‌شود.

(FP): مثبت کاذب، اقلامی که برای کاربران جذاب نیست ولی به آن‌ها توصیه می‌شود.

(FN): اقلامی که توصیه نمی‌شوند اما کاربران به آن‌ها علاقه دارند.

(TN): اقلامی که به‌درستی توصیه نشدند و موردعلاقه کاربر نیستند.

$$Recall = \frac{1}{n}\sum_{i=1}^{n}\frac{TP}{TP+FN} \quad ۲۰$$

$$F1 = \frac{2\times precision \times recall}{precision + recall} \quad ۲۱$$

$$MAE = \frac{\sum_{i=1}^{N}|r_i - p_i|}{N} \quad ۲۲$$

که p_i و r_i به ترتیب امتیاز پیش‌بینی‌شده و واقعی قلم i و N تعداد کل امتیازات پیش‌بینی‌شده توسط سیستم پیشنهادی است.

$$RC = \frac{\#number\ of\ \ predicted\ ratings}{\#number\ of\ \ all\ ratings} \quad ۲۳$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(r_i - p_i)} \quad ۲۴$$

روش پیشنهادی بر روی مجموعه داده Epinions مورد ارزیابی قرار گرفته است. جداول ۱، ۲ و ۳، مقادیر معیارهای Precision، Recall و F1 الگوریتم‌های مختلف را به ترتیب برای مقادیر مختلف N: ۵، ۱۰، ۱۵، ۲۰ و ۳۰ برای توصیه کا-بالاترین قلم نشان می‌دهد. همان‌طور که در جدول ۱ نشان داده شده است، روشTRACCF بهترین عملکرد را برای تمام مقادیرN به‌جز مقدار ۳۰ دارد و Precision با افزایش مقدار کا کاهش می‌یابد.

باوجودآنکه روش پیشنهادی برای معیار دقت از روش TRCCF برای تعداد کمتر اقلام پیشنهادی بهتر عمل نکرده است اما نتایج معیار-اف بیانگر عملکرد بهتر روش پیشنهادی، نسبت به این الگوریتم و کلیه الگوریتم‌ها است. معیار–اف تعادلی میان معیار دقت و فراخوان هست. فقط این نکته لازم به ذکر است که با افزایش تعداد اقلام توصیه داده‌شده، دقت کارایی روش پیشنهادی نیز نسبت به الگوریتم TRACCF افزایش می‌یابد. در معیار دقت چون FP ها در نظرگرفته نمی‌شوند و ازآنجایی‌که مقدار این در سیستم پیشنهادی بیشتر است نتایج دقت روش‌های دیگر از روش پیشنهادی بهتر شده است. به همین منظور معیار-اف را محققین به‌عنوان نتیجه ارزیابی سیستم که مبتنی بر FNهای سیستم هم هست را معیار مناسب سنجش در نظر می‌گیرند.

جدول ۱. مقایسه الگوریتم‌ها از جنبه معیار Precision

Algorithm	Top5	Top10	Top15	Top20	Top30
TRACCF	0.9462	0.8981	0.8981	0.8981	0.5319
KMCF-U	0.7857	0.7857	0.7857	0.7857	0.7367
KMCF-I	0.7472	0.7361	0.7361	0.7361	0.7361
TOTAR(Feng, Tian, Wang, & Li, 2015)	0.8550	0.8521	0.8416	0.8281	0.8237
روش پیشنهادی	0.8502	0.8527	0.8543	0.8541	0.8543

جدول ۲. مقایسه الگوریتم‌ها از جنبه معیار Recall

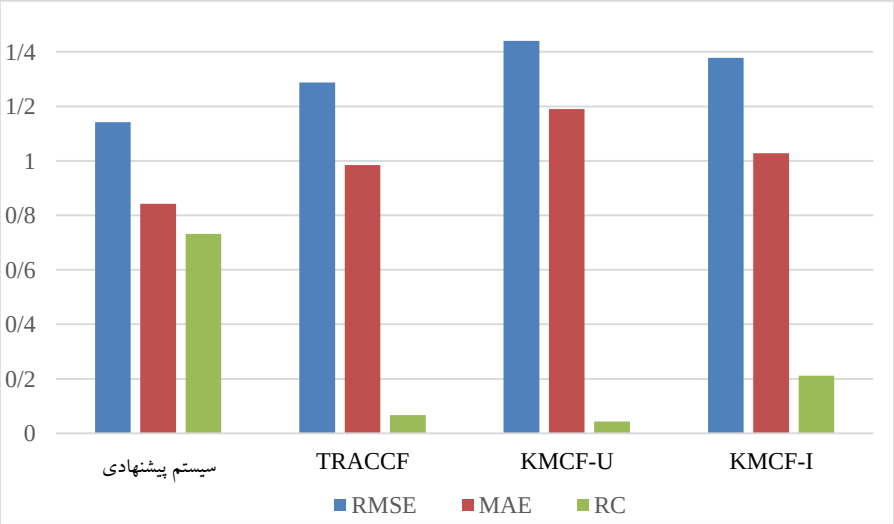
Algorithm	Top5	Top10	Top15	Top20	Top30
TRACCF	0.5239	0.5399	0.5399	0.5399	0.5481
KMCF-U	0.4158	0.4158	0.4158	0.4158	0.4257
KMCF-I	0.6481	0.6481	0.6481	0.6481	0.6481
TOTAR (Feng et al., 2015)	0.4356	0.5467	0.7322	0.7985	0.8134
روش پیشنهادی	0.8152	0.8509	0.8635	0.8657	0.8679

جدول ۳ مقایسه الگوریتم‌ها از جنبه معیار F1-measure

Algorithm	Top5	Top10	Top15	Top20	Top30
TRACCF	0.6744	0.6744	0.6744	0.6744	0.5399
KMCF-U	0.5438	0.5438	0.5438	0.5438	0.5395
KMCF-I	0.6941	0.6893	0.6893	0.6893	0.6693
TOTAR (Feng et al., 2015)	0.5636	0.6660	0.7830	0.8130	0.8185
روش پیشنهادی	0.8323	0.8518	0.8588	0.8599	0.8610

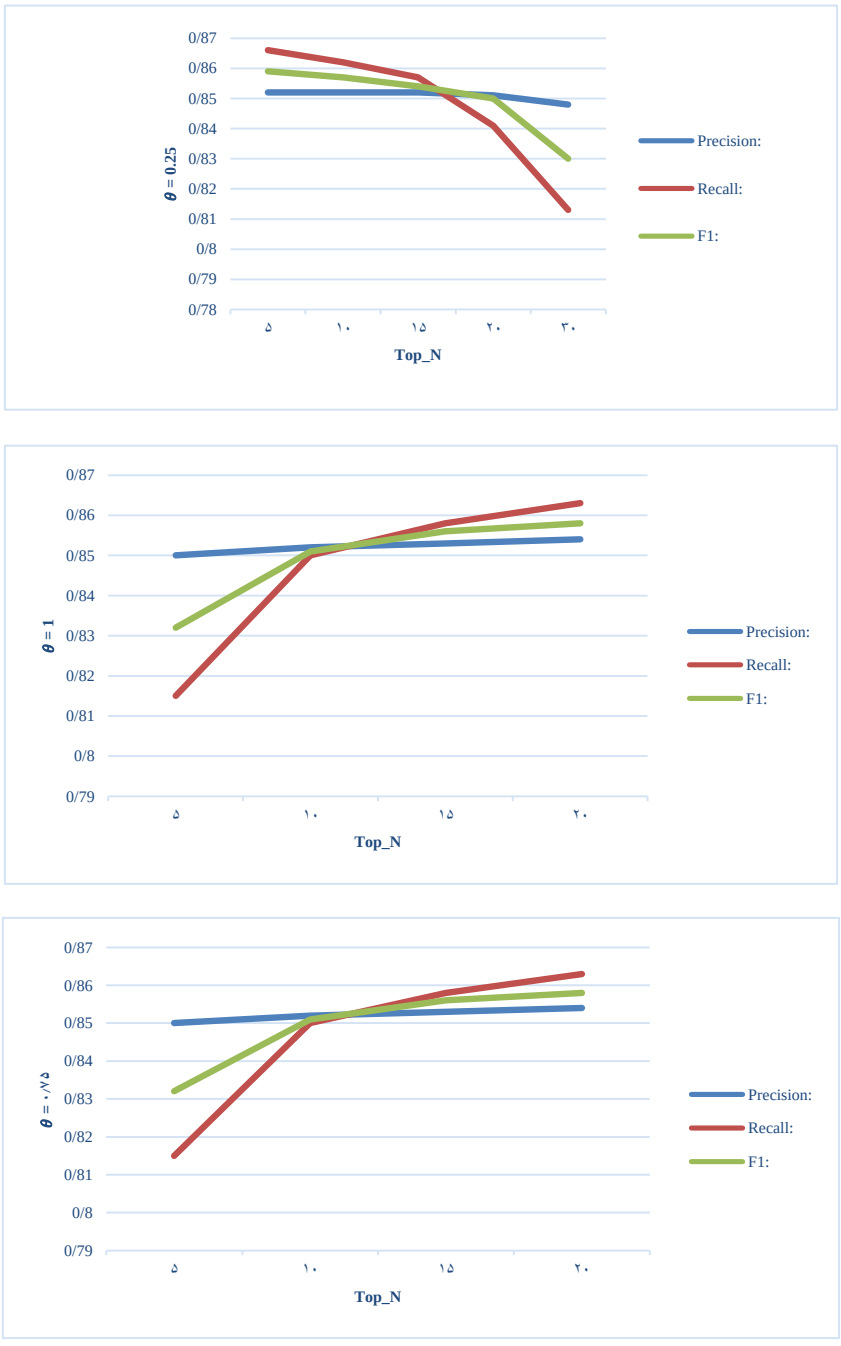
شکل ۳ نتایج مقایسه سیستم توصیه‌گر پیشنهادی با روش‌های دیگر در سه معیار MAE, RMSE و RC را مشاهده می‌نمایید. همانطور که مشاهده می‌نمایید روش پیشنهادی در مقایسه با سایر روش‌های مشابه خطای کمتری دارد.

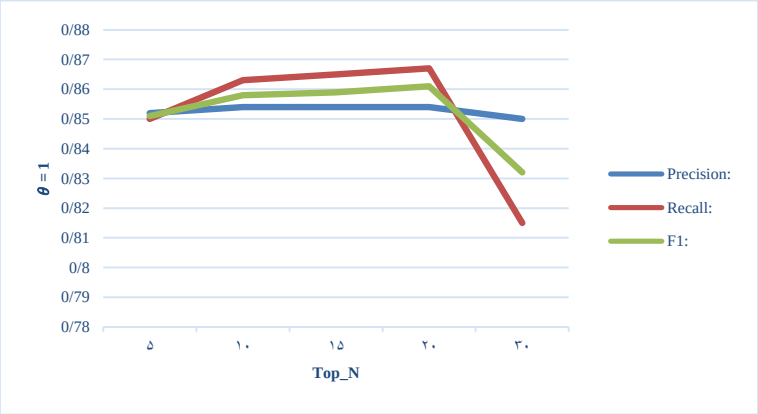
شکل ۳. نمودار مقایسه الگوریتم‌ها از جنبه معیارهای ارزیابی MAE, RMSE,RC

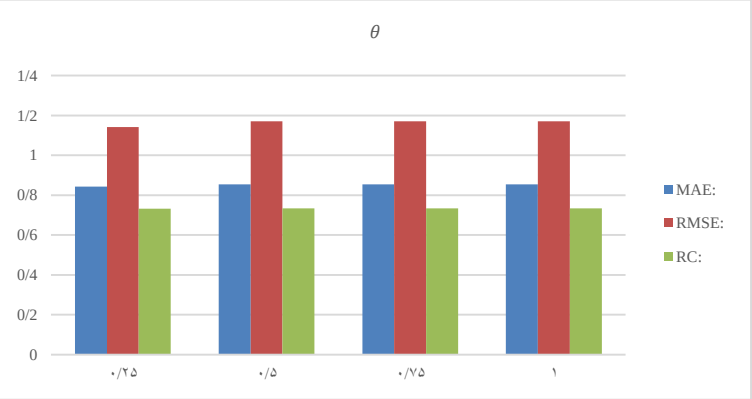


روش پیشنهادی شامل دو پارامتر قابل تنظیم (φو θ) است و عملکرد آن به تخصیص مقادیر مناسب به این پارامترها بستگی دارد. شکل ۴ و ۵ نتایج تجربی را با در نظر گرفتن مقادیر مختلف این پارامترها نشان می‌دهند. با توجه به این نتایج، بهترین مقدار برای مقایسه روش پیشنهادی با سایر الگوریتم‌ها $\theta = 0.25$ و $\varphi = 1$ در نظر گرفته شده است.

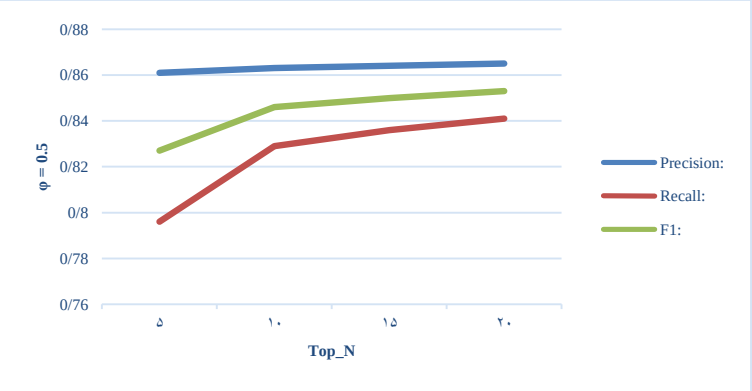
شکل ٤. تأثیر پارامتر θ بر روی روش پیشنهادی

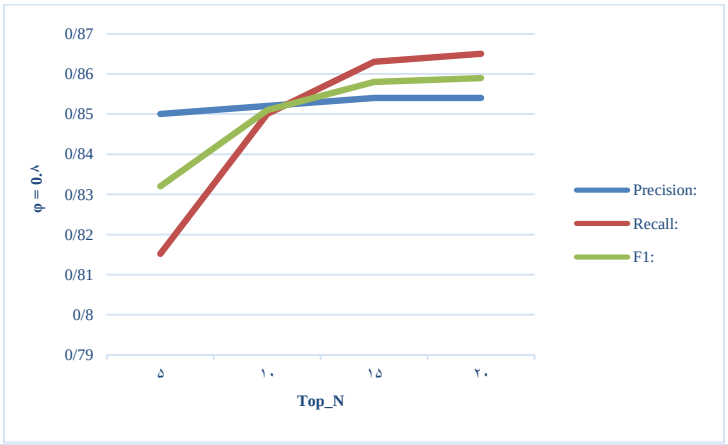


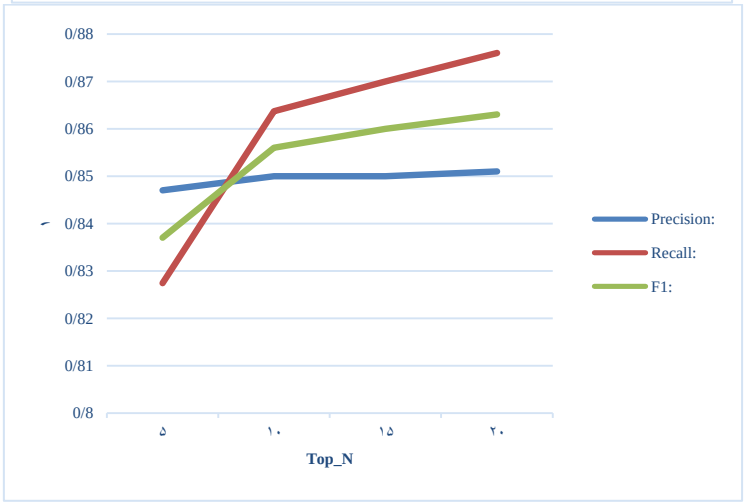


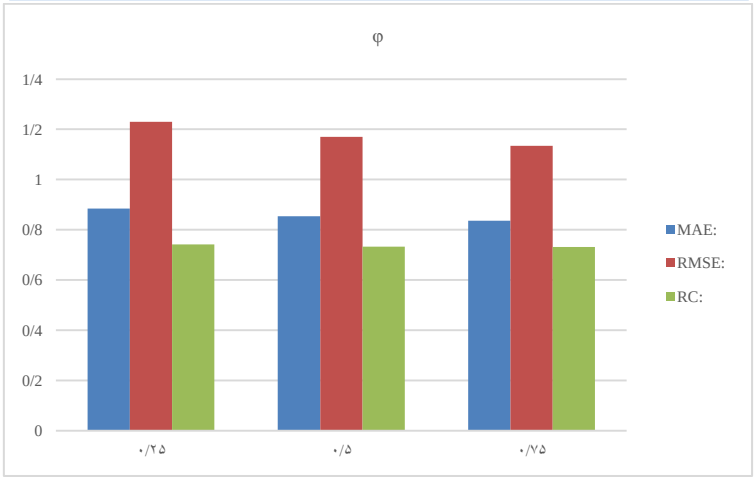


شکل ۵. تأثیر پارامتر φ بر روی روش پیشنهادی.




0/87
0/86
0/85
0/84
0/83
0/82
0/81
0/8
0/79
φ = 0.۸
۵
۱۰
۱۵
۲۰
Top_N
Precision:
Recall:
F1:


0/88
0/87
0/86
0/85
0/84
0/83
0/82
0/81
0/8
۵
۱۰
۱۵
۲۰
Top_N
Precision:
Recall:
F1:


φ
1/4
1/2
1
0/8
0/6
0/4
0/2
0
۰/۲۵
۰/۵
۰/۷۵
MAE:
RMSE:
RC:

در جدول ۴ زمان موردنیاز برای اجرای الگوریتم‌ها نمایش داده شده است. تعداد کاربران برابر n و تعداد اقلام برابر m و c نشان‌دهنده تعداد خوشه‌ها است.

جدول ٤. پیچیدگی زمانی روش پیشنهادی در مقایسه با سایر روش‌ها.

Algorithm	Time	Trust	Graph-based	Time complexity
TRACCF	No	Yes	No	$O(nmc + nm)$
KMCF-U	No	No	No	$O(nmc + nm)$
KMCF-I	No	No	No	$O(nmc + nm)$
TOTAR (Feng et al., 2015)	Yes	No	Yes	$O(n^2m)$
روش پیشنهادی	Yes	Yes	Yes	$O(n^2m)$

نتیجه‌گیری

امروزه سیستم‌های توصیه‌گر کاربردهای بسیاری در خریدهای بر خط، شبکه‌های اجتماعی و فیلم و موزیک و غیره داشته و از اطلاعات منابع مختلف مانند داده‌های متنی کاربران و اقلام، امتیازات (یا خرید)، تاریخچه تراکنش کاربران برای ارائه توصیه‌های منطبق با سلیقه کاربران استفاده می‌کنند. در بسیاری از روش‌های بکار گرفته‌شده، زمان امتیازدهی اقلام یا تراکنش‌ها در نظر گرفته نمی‌شود. درصورتی‌که ممکن است سلیقه کاربر در دوره‌های زمانی مختلف تغییر کند. همچنین وجود افراد جعلی در شبکه‌های اجتماعی سبب کاهش دقت توصیه‌های سیستم‌های توصیه‌گر شده است. در این مقاله، یک سیستم توصیه‌گر بر اساس زمان امتیازدهی و همچنین روابط اعتماد بین کاربران با استفاده از یک الگوریتم تشخیص جوامع همپوشان جدید برای کشف ساختار جامعه بین کاربران پیشنهاد شده است. الگوریتم پیشنهادی پراکندگی ماتریس کاربر_قلم را با افزودن امتیازات ضمنی و حل مشکل شروع سرد برای کاربران جدید را با در نظر گرفتن اعتماد بین کاربران حل نموده است. نتایج ارزیابی سیستم پیشنهادی در قالب یک سیستم توصیه‌گر مبتنی بر فیلم بر روی مجموعه داده‌ای Epinions نشانگر کارایی بالاتر سیستم پیشنهادی نسبت به سایر سیستم‌های مرسوم و پایه مانند TRACCF، KMCF-U، KMCF-I، TOTAR برای ارائه کا-بالاترین قلم بدون افزایش پیچیدگی زمانی می‌باشد. جهت کارهای آتی روش پیشنهادی را جهت شناسایی حملات در سیستم‌های توصیه‌گر پالایش همکارانه می‌توان

موردبررسی قرار داده و جوامع جعلی را شناسایی نموده و تأثیر آن‌ها بر روی عملکرد سیستم را کاهش دهیم.

ORCID

Fatemeh Rezaimehr **https://orcid.org/0000-0001-7579-8094**

Chitra Dadkhah **https://orcid.org/0000-0002-9836-9388**

References

Abdul-Rahman, A., & Hailes, S. (2000). *Supporting trust in virtual communities.* Paper presented at the 33th Hawaii International Conference on System Sciences, Hawaii, USA.

Afsar, M. M., Crump, T., & Far, B. (2022). Reinforcement learning based recommender systems: A survey. *ACM Computing Surveys, 55*(7), 1-38.

Aggarwal, C. C. (2016). *Recommender systems* (Vol. 1): Springer.

Ahmadian, S., Joorabloo, N., Jalili, M., Ren, Y., Meghdadi, M., & Afsharchi, M. (2020). A social recommender system based on reliable implicit relationships. *Knowledge-Based Systems, 192*, 105371.

Alahmadi, D. H., & Zeng, X.-J. (2015). ISTS: Implicit social trust and sentiment based approach to recommender systems. *Expert Systems with Applications, 42*(22), 8840-8849. doi:http://dx.doi.org/10.1016/j.eswa.2015.07.036

Bansal, J., Flannery, M. B., & Woolhouse, M. H. (2021). Influence of personality on music-genre exclusivity. *Psychology of Music, 49*(5), 1356-1371.

Casillo, M., Gupta, B. B., Lombardi, M., Lorusso, A., Santaniello, D., & Valentino, C. (2022). Context aware recommender systems: A novel approach based on matrix factorization and contextual bias. *Electronics, 11*(7), 1003.

Cho, J., Kwon, K., & Park, Y. (2009). Q-rater: a collaborative reputation system based on source credibility theory. *Expert Systems with Applications, 36*, 3751-3760.

Daneshmand, S. H., Javari, A., Abtahi, S. E., & Jalili, M. (2015). A time-aware recommender system based on dependency network of items. *The Computer Journal, 58*(9), 1255-1266.

Feng, H., Tian, J., Wang, H. J., & Li, M. (2015). Personalized recommendations based on time-weighted overlapping community detection. *Information & Management, 52*(7), 789-800.

G. Guo, J. Zhang, & Thalmann, D. (2014). Merging trust in collaborative filtering to alleviate data sparsity and cold start. *Knowledge-Based Systems, 57*, 57-68.

Gharibshah, J., & Jalili, M. (2014). Connectedness of users–items networks and recommender systems. *Applied Mathematics and Computation, 243*, 578-584.

Ghavipour, M., & Meybodi, M. R. (2016). An adaptive fuzzy recommender system based on learning automata. *Electronic Commerce Research and Applications, 20*, 105-115. doi:http://dx.doi.org/10.1016/j.elerap.2016.10.002

Guo, G., Zhang, J., & Thalmann, D. (2014). Merging trust in collaborative filtering to alleviate data sparsity and cold start. *Knowledge-Based Systems, 57*, 57-68.

Hernando, A., Bobadilla, J., Ortega, F., & Tejedor, J. (2013). Incorporating

reliability measurements into the predictions of a recommender system. *Information Sciences, 218*, 1-16.

Hui, B., Zhang, L., Zhou, X., Wen, X., & Nian, Y. (2022). Personalized recommendation system based on knowledge embedding and historical behavior. *Applied Intelligence, 52*(1), 954-966.

Ingoo, H., Kyong, J. O., & Tae, H. R. (2003). The collaborative filtering recommendation based on SOM cluster-indexing CBR. *Expert Systems with Applications, 25*, 413–423.

Jiang, M., Cui, P., Wang, F., Zhu, W., & Yang, S. (2014). Scalable Recommendation with Social Contextual Information. *IEEE Transactions on Knowledge and Data Engineering, 26*(11), 2789-2802. doi:10.1109/TKDE.2014.2300487

Koren, Y. (2010). Collaborative filtering with temporal dynamics. *Communications of the ACM, 53*(4), 89-97.

Lathia, N., Hailes, S., & Capra, L. (2008). *Trust-based collaborative filtering.* Paper presented at the Proceedings of trust management II.

Liu, F., & Lee, H. (2010). Use of social network information to enhance collaborative filtering performance. *Expert Systems with Applications, 37*, 4772–4778.

Manolopoulus, Y., Nanopoulus, A., Papadopoulus, A. N., & Symeonidis, P. (2008). Collaborative recommender systems: combining effectiveness and efficiency. *Expert Systems with Applications, 34*, 2995-3013.

Massa, P., & Avesani, P. (2007). *Trust-aware recommender systems.* Paper presented at the Proceedings of the 2007 ACM conference on Recommender systems.

Moradi, P., & Ahmadian, S. (2015). A reliability-based recommendation method to improve trust-aware recommender systems. *Expert Systems with Applications, 42*(21), 7386-7398.

Moradi, P., Ahmadian, S., & Akhlaghian, F. (2015). An effective trust-based recommendation method using a novel graph clustering algorithm. *Physica A: Statistical Mechanics and its Applications, 436*, 462-481. doi:http://dx.doi.org/10.1016/j.physa.2015.05.008

P. Massa, & Avesani, P. (2007). *Trust-aware recommender systems.* Paper presented at the 2007 ACM conference on Recommender systems, Minneapolis, Minnesota, USA.

P. Moradi, & Ahmadian, S. (2015). A reliability-based recommendation method to improve trust-aware recommender systems. *Expert Systems with Applications, 42*(21), 7386-7398. doi:http://dx.doi.org/10.1016/j.eswa.2015.05.027

Ranjbar Kermany, N., & Alizadeh, S. H. (2017). A hybrid multi-criteria recommender system using ontology and neuro-fuzzy techniques. *Electronic Commerce Research and Applications, 21*, 50-64. doi:http://dx.doi.org/10.1016/j.elerap.2016.12.005

Rezaimehr, F., & Dadkhah, C. (2021). A survey of attack detection approaches in collaborative filtering recommender systems. *Artificial*

Intelligence Review, 54(3), 2011-2066.
Ricci, F. (2010). Mobile recommender systems. *Information Technology & Tourism, 12*(3), 205-231. doi:https://doi.org/10.3727/109830511X12978702284390
S. Ahmadian, M. Meghdadi, & Afsharchi, M. (2018a). Incorporating reliable virtual ratings into social recommendation systems. *Applied Intelligence, 48*(11), 4448-4469.
S. Ahmadian, M. Meghdadi, & Afsharchi, M. (2018b). A social recommendation method based on an adaptive neighbor selection mechanism. *Information Processing & Management, 54*(4), 707-725.
S. Ahmadian, P. Moradi, & Akhlaghian, F. (2014). *An improved model of trust-aware recommender systems using reliability measurements.* Paper presented at the 2014 6th Conference on Information and Knowledge Technology (IKT).
Tohidi, N., & Dadkhah, C. (2020). Improving the performance of video Collaborative Filtering Recommender Systems using Optimization Algorithm. *International Journal of Nonlinear Analysis and Applications, 11*(1), 283-295.
Wu, S., Sun, F., Zhang, W., Xie, X., & Cui, B. (2022). Graph neural networks in recommender systems: a survey. *ACM Computing Surveys, 55*(5), 1-37.
X. Yang, Y. Guo, & Liu, Y. (2013). Bayesian-Inference-Based Recommendation in Online Social Networks. *IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, 24*, 642-651.
Xiong, L., Chen, X., Huang, T.-K., Schneider, J., & Carbonell, J. G. (2010). Temporal Collaborative Filtering with Bayesian Probabilistic Tensor Factorization. In *Proceedings of the 2010 SIAM International Conference on Data Mining* (pp. 211-222).
Yang, X., Guo, Y., & Liu, Y. (2012). Bayesian-inference-based recommendation in online social networks. *IEEE Transactions on Parallel and Distributed Systems, 24*(4), 642-651.

استناد به این مقاله: رضایی‌مهر، فاطمه.، دادخواه، چیترا. (۱۴۰۲). سیستم توصیه‌گر زمان و اعتماد مبنا با استفاده از تشخیص جوامع مبتنی بر گراف، *مطالعات مدیریت کسب وکار هوشمند*، ۱۲(۴۶)، ۳۲۷-۳۶۰.
DOI: 10.22054/ims.2023.70722.2247

Journal of Business Intelligence Management Studies is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License..