Expert Systems With Applications 238 (2024) 121887

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

IUG-CF: Neural collaborative filtering with ideal user group labels

Zi-Feng Peng, Heng-Ru Zhang *, Fan Min

School of Computer Science, Southwest Petroleum University, Chengdu, 610500, China
Lab of Machine Learning, Southwest Petroleum University, Chengdu, 610500, China

ARTICLE INFO

Keywords:
Collaborative filtering
Demographic information
Ideal user group
Neural networks
Recommender systems

ABSTRACT

Demographics are crucial information for recommender systems (RSs). Most existing demographic-based RSs focus on similarity between user profiles. However, they rarely incorporate demographic data to describe an item and establish the connection between items and users. In this paper, we propose the concept of the ideal user group (IUG) as a dynamic label for items. This label indicates the users who are most suitable for an item, based on the demographics of its historical customers. Unlike a general label (such as genre or language), the IUG is dynamically changing with the distribution of historical user demographics and is built based on demographic information that undergoes a split-combine process. To validate our method's effectiveness, we propose an IUG-based neural collaborative filtering (IUG-CF) model. Experimental results on three real-world datasets show that the IUG is an effective approach for improving recommendation performance.

1. Introduction

In the era of information overload, recommender systems (RSs) play an important role in providing more suitable movies, music, books, etc. There are various types of RSs, including content-based RS, demographic-based RS (DRS), collaborative filtering (CF), and hybrid RS (Al-Shamri, 2016; Bobadilla, Ortega, Hernando, & Gutiérrez, 2013). Content-based RS makes recommendations based on historical rating records and extracted text or image information (Van Meteren & Van Someren, 2000). DRS connects users through demographic similarity, and its rationality is based on similar user groups having similar preferences (Krulwich, 1997; Pazzani, 1999; Rich, 1979). CF explores collaborative information through user–item interactions and user or item similarities (Khojamli & Razmara, 2021; Konstan et al., 1997; Resnick, Iacovou, Suchak, Bergstrom, & Riedl, 1994; Shen, Zhang, Yu, & Min, 2019; Zhang, Min, Zhang, & Wang, 2019). Hybrid RS combines user–item interaction information with other information such as demographics and text (Sarker & Matin, 2021; Vozalis & Margaritis, 2007).

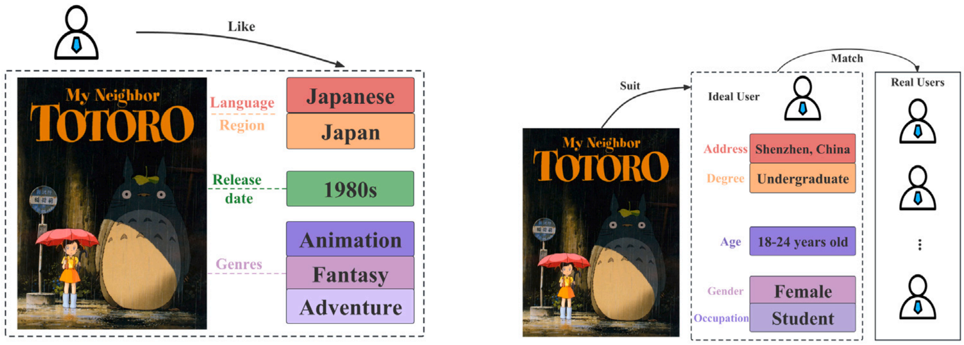
People who share certain demographics are likely to have similar preferences, which is the underlying principle of DRS (Chen, Wu, et al., 2023; Krulwich, 1997; Pazzani, 1999; Rich, 1979; Tiwari, Ashpilaya, Vedita, Daripa, & Paltani, 2020). While demographic information, such as age, gender, and occupation, can be helpful in making personalized recommendations, it is important to recognize that relying solely on demographics can lead to biased recommendations and perpetuate stereotypes. Most existing approaches focus on leveraging the relationship between a user and item labels to make predictions. This prediction is often based on the assumption that if the user has liked items with similar labels in the past, they are likely to like the current item (Mooney & Roy, 2000; Wang, Liang, Xu, Feng, & Guan, 2018), as depicted in Fig. 1(a). In contrast, we adopt a different approach where we use user information to describe items. By doing so, we transform the problem of predicting whether a user likes a current item into the task of determining whether the item is suitable for the specific user, as depicted in Fig. 1(b). Our approach involves constructing an ideal user group (IUG) that corresponds to the item and then determining user behavior based on their degree of matching to this user group.

In this paper, we investigate the popularity of items across different demographic attributes and combine individual demographic attributes that show stronger preferences. This process involves splitting and combining demographic attributes, defining the resulting combined attributes as an IUG. To construct an IUG, we first separate user groups based on each demographic attribute, such as gender, occupation, etc. Next, we calculate the average rating of each group based on historical rating records and employ Bayesian averaging to address the issue of certain user groups having fewer members but stronger preferences (Yang & Zhang, 2013). We then compare the popularity of each item across different demographics and combine the demographic values corresponding to the group with the maximum average rating of each attribute to obtain the demographics of an IUG. We highlight that the IUG is both "built-up" and "dynamic" as illustrated in Fig. 2.

To validate the effectiveness of our recommendation process, we present the IUG-CF model which combines our proposed IUG concept

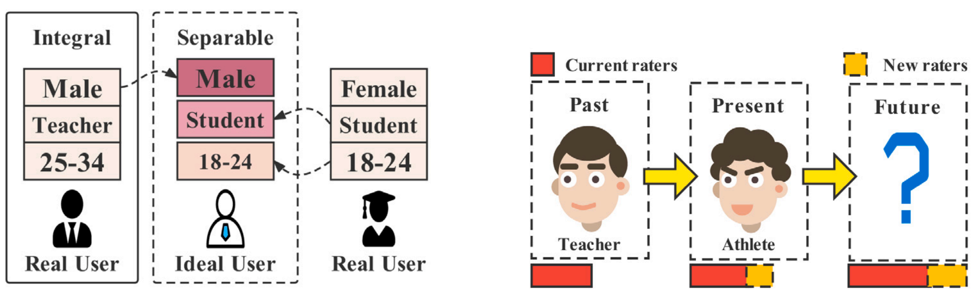
* Corresponding author at: School of Computer Science, Southwest Petroleum University, Chengdu, 610500, China.
E-mail addresses: pocket_eight@163.com (Z.-F. Peng), zhanghr@swpu.edu.cn (H.-R. Zhang), minfan@swpu.edu.cn (F. Min).

https://doi.org/10.1016/j.eswa.2023.121887
Received 9 April 2023; Received in revised form 6 September 2023; Accepted 27 September 2023
Available online 6 October 2023



(a) Common method (b) Our method

Fig. 1. Comparison of our method with commonly used methods in describing and recommending movies. (a) The movie is described by genre, language, and other information. (b) The movie is described as best suited for a user group with certain demographics.



(a) Built-up (b) Dynamic

Fig. 2. A demonstration of two characteristics of the IUG. (a) The demographics of an IUG are merged by the demographics of two kinds of user groups. (b) The IUG is dynamically changing with the item's raters increasing and the width of the rectangle representing the number of raters.

with a neural network. The IUG-CF model has two parts. The first part involves constructing IUGs for all items to calculate the preference degree for each demographic value. The second part involves generating the recommendation list by mapping user demographics to vectors and calculating the distance between them. In order to match users with IUGs of items, the demographics are mapped as vectors and the distance between them is used. Fig. 3 is a toy example of the recommended process.

We summarize the contributions as follows:

- We propose the concept of IUG, which, to the best of our knowledge, is a novel approach in DRS. By defining items using IUG, we offer a fresh perspective that avoids the cold-start problem.
- We propose a new neural collaborative filtering model called IUG-CF, which utilizes IUGs to improve recommendation performance. The model takes into account item popularity across different demographic categories and recommends items to users with similar demographic profiles to the IUGs of the items.
- We introduce a new technique for connecting users and items by splitting and combining demographic information. This method enables the DRS to fully utilize demographic information to provide better recommendations.

The remainder of this paper is structured as follows. Section 2 introduces related work. Section 3 details our IUG-CF model. In Section 4, we report our results and discuss the effects of various parameters. Finally, Section 5 draws the conclusion of this paper.

2. Related works

In this section, we present a review of the relevant literature in the field. Section 2.1 focuses on the prior research related to demographic-based recommendations, while Section 2.2 delves into the previous works concerning Deep-Learning based recommendations.

2.1. Demographic-based recommendations

The DRS is a traditional model that considers users' demographic information, such as age, gender, occupation, etc., as an essential factor for personalized recommendations. The concept of modeling users' interests based on demographic information was first proposed by Grudy in 1979 (Rich, 1979), which laid the foundation for exploring the role of demographics in RS. In the 1990s, DRSs became more prevalent, such as the Lifestyle Finder (Krulwich, 1997), which used user demographic data to create user profiles and recommend web pages. Pazzani proposed a model that utilized demographic information from user homepages to calculate similarity among users (Pazzani, 1999). This model inspired the development of various similarity calculation metrics between user profiles (Sahu & Dwivedi, 2019; Vozalis & Margaritis, 2007).

During the first decade of the 21st century, researchers attempted to integrate demographic information into hybrid RSs (Burke, 2002; Lekakos & Giaglis, 2007). In recent years, demographic information has been increasingly used in recommendation algorithms for e-commerce (Souali, El Afia, & Faizi, 2011) and social platforms (Zhao et al., 2014, 2016). Moreover, researchers have explored various ways to

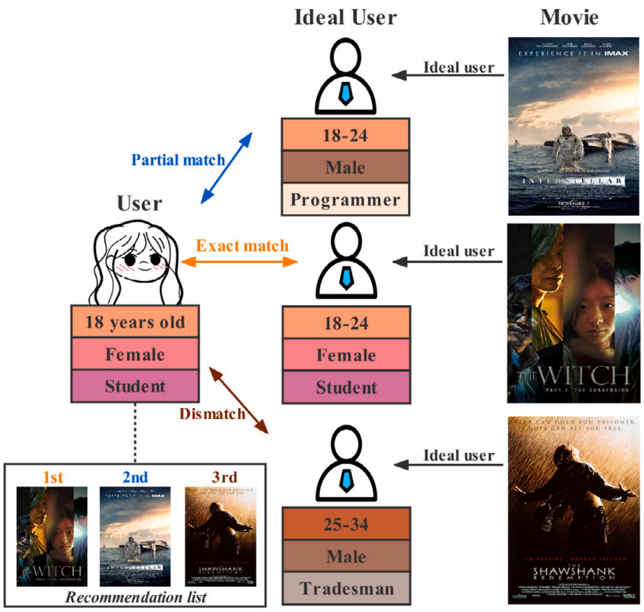


Fig. 3. A toy example of generating recommendation list for a user.

incorporate demographic information into recommendation systems. For example, Tahmasebi, Meghdadi, Ahmadian, and Valiallahi (2021) combined cosine similarity and demographic similarity to enhance the neighborhood set of users, while Yassine, Mohamed, and Al Achhab (2021) proposed a model based on item feature selection that incorporates demographic information to segment users. Bobadilla, González-Prieto, Ortega, and Lara-Cabrera (2021) demonstrated the importance of feature extraction of demographic information in addition to item feature extraction, and Liu, Qu, Chen, and Mahmud (2019) used rating records to infer user demographic information, which mitigated the difficulty of obtaining such information.

Furthermore, group recommender systems focus on recommending items to a group of people (Sacharidis, 2019). Both explaining recommendation and aggregation strategies are crucial subtasks within the domain of group recommendation (Jameson, 2004; Masthoff, 2010, 2015). Demographic information is frequently incorporated in these subtasks to deliver personalized recommendations for users (Renjith, Sreekumar, & Jathavedan, 2020).

The cold-start problem is a well-known issue with RSs. Many researchers have shown that DRS is a popular choice when attempting to solve the cold-start problem (Pandey & Rajpoot, 2016; Safoury & Salah, 2013; Tahmasebi et al., 2021). Common methods used in DRS include calculating the similarity between new users and existing users (Al-Shamri & Bharadwaj, 2007; Tahmasebi et al., 2021) and co-clustering (Pereira & Hruschka, 2015).

Traditional DRSs mainly focus on segmenting users based on demographic information and exploring how different similarity calculation metrics can improve recommendation performance. Building on previous work, our proposed IUG-CF model maps demographic attributes into a high-dimensional space, linking each item with ideal demographics that are combined by certain demographics.

2.2. Deep-learning based recommendations

In recent years, DL-based RSs have gained popularity due to their impressive performance (Covington, Adams, & Sargin, 2016; Kim, Park, Oh, Lee, & Yu, 2016; Xue et al., 2019). Deep Learning algorithms are employed to learn the intricate mapping relationship between user and item through latent factor representation and matching functions. NeuMF (He et al., 2017) is a widely used model that combines MF and MLP to capture both linear and nonlinear interactions between users and items. By merging these two approaches, it gains the ability to learn arbitrary functions from the given data. In recent years, several studies have been conducted based on the NeuMF model. For instance, NHF (Sarker & Matin, 2021) adds auxiliary information of items and users to the model, such as demographic information, to further understand the interaction between items and users. HybridNNMF (Zeng et al., 2021) combines a global multi-layer perceptron and local residual networks to model user–item interactions effectively. Furthermore, in DL-based recommendation models, the attention mechanism is commonly integrated. For example, LRML (Tay, Anh Tuan, & Hui, 2018) adopts a memory-based attention network to represent latent user–item relationships as latent relation vectors. Similarly, IFDNAMF (Tian, Pan, Yin, & Wang, 2021) incorporates attribute information and utilizes element-wise product for cross-feature learning during information fusion, leveraging the attention mechanism to discern the importance of different cross-features in prediction outcomes. Indeed, some models focus on the exploration and utilization of item and user features. For instance, CFFNN (Yu et al., 2021) utilizes a cross feature fusion network to enhance collaborative filtering by combining user features and item features. The Wide&Deep model (Cheng et al., 2016), on the other hand, employs a "wide" component with a linear model to capture explicit interactions among features, while utilizing a "deep" component with a DNN to learn hierarchical feature representations. Additionally, CoupledCF (Zhang, Cao, Zhu, Li, & Sun, 2018) is a model that simultaneously learns both explicit and implicit couplings within and between users and items concerning attributes and deep features.

It is worth noting that demographic information is extensively utilized in DL-based recommendation models. For example, the four models mentioned above, Wide&Deep, NHF, IFDNAMF, and CoupleCF, all employ demographic information to assist the model learn more user characteristics. By incorporating demographic information, DL-based RSs can better understand user needs, leading to an enhanced user experience and improved recommendation performance.

Table 1
Notations.

Notation	Meaning
$\mathcal{I}, \mathcal{J}$	the set of users and items
$\mathcal{I}_j$	the set of users who have rated item j
$\mathcal{I}_j^p$	the set of users with value p who have rated item j
$\mathcal{J}_i$	the set of items rated by user i
$\mathcal{P}$	the set of all demographic values
$\mathcal{P}_q$	the set of demographic values of attribute $q \in \mathcal{Q}$
$\mathcal{P}_j$	the set of demographic values that belong to the ideal user group of j
$\mathcal{Q}$	the set of demographic information attributes
$\mathcal{R}$	the set of user–item interactions
$\mathcal{U}'$	the set of all IUs
$\mathbf{p}_j$	the embedding for item $j \in \mathcal{J}$
$\mathbf{q}_i$	the embedding for user $i \in \mathcal{I}$
$\mathbf{d}^p$	the embedding for value $p \in \mathcal{P}$
c_j	the average rating of item j
c_j^p	the average rating of item j on demographic value $p \in \mathcal{Q}$
r_{ij}	the user i's rating of item j
s_{ij}	the similarity of user i and the IU of item j
u_i	the demographics of user i
u'_j	the user member of item j's IUG

3. The method

In this section, we introduce the problem statement and the overall flow of the proposed IUG-CF model. The model consists of two main components: building IUGs for all items using historical user–item interactions and generating personalized recommendations based on demographic information matching. The flow of the model is illustrated in Fig. 4.

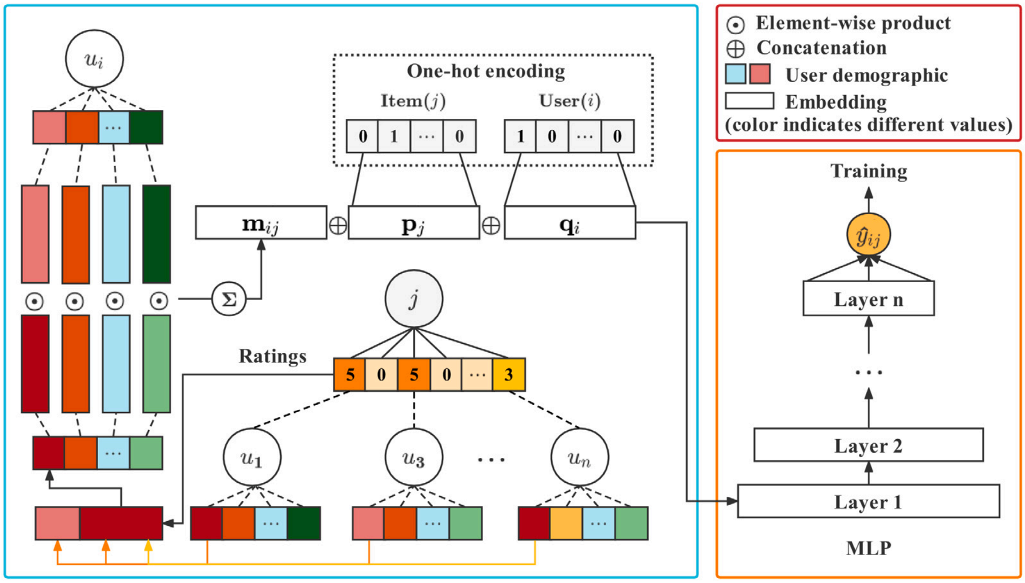


Fig. 4. Overview of the IUG-CF model.

3.1. Problem statement

Table 1 lists the notations used in the paper.

In the MovieLens-1M dataset (Harper & Konstan, 2015), demographic information includes attributes such as age, gender, and occupation. Each attribute has specific values, such as male and female for gender. Therefore, we have the following two definitions regarding demographic information and its values.

Remark 1. Demographics attributes and their corresponding values are defined as $q \in \mathcal{Q}$ and $p_q \in \mathcal{P}_q$ respectively. $\forall q_g, q_h \in \mathcal{Q}$, s.t. $\mathcal{P}_{q_g} \bigcap \mathcal{P}_{q_h} = \emptyset$, and $\mathcal{P}_{q_1} \bigcup \mathcal{P}_{q_2} \bigcup \cdots \bigcup \mathcal{P}_{q_n} = \mathcal{P}$.

Remark 2. User i is defined as a set $u_i = p_1, p_2, \ldots$, where p denotes the demographic value.

For example, $u_i = \{p_{q_1}, p_{q_2}, p_{q_3}\}$, which means the user has the demographic values p_{q_1}, p_{q_2}, and p_{q_3}. Here, q_1, q_2, and q_3 represent different demographic attributes, such as age, gender, and occupation, and p_{q_k} denotes the corresponding value of the kth attribute for the user.

3.2. Ideal user building

The first step of IUG-CF involves building the IUG. For each project, our target is to establish a precise one-to-one mapping between the item and the IUG. Each user group possesses its own demographic information. To construct this IUG, we initially examine the preferences and likings of items based on different demographic attributes. While the average rating is commonly used to measure item popularity, the demographic distribution of item consumers is often non-uniform. To address this issue, we employ Bayesian averaging (Raftery, Madigan, & Hoeting, 1997), a technique used on platforms such as IMDB[1] to ensure a uniform minimum number of rating records for each demographic. This results in more representative and fair average ratings for each demographic. Specifically, the average rating of item j over the demographic value $p \in \mathcal{P}_q$ is expressed as

$$c_j^p = \frac{a_j^q \times c_j + \sum_{i \in \mathcal{I}_j} r_{ij} \cdot B(p \in u_i)}{|\mathcal{I}_p^j| + a_j^q}, \tag{1}$$

where $a_j^q = \gamma \frac{|\mathcal{I}_j|}{\mathcal{P}_q}$ is the extended number of rating records; γ is a coefficient that regulates the amount of the extension; c_j is the average rating of all items; $\mathcal{I}_j^p$ is the set of users who have demographic value p and have rated item j; the function $B(\cdot)$ is defined as

$$B(x) = \begin{cases} 1 & (x = \text{True}); \\ 0 & (x = \text{False}). \end{cases} \tag{2}$$

To evaluate the popularity of an item across different user groups, we treat each demographic attribute as a distinct dimension. Specifically, we use the set C_j to represent the popularity of item j across different demographic attributes, denoted as

$$C_j = \{\mathbf{c}_{(j,q_1)}, \mathbf{c}_{(j,q_2)}, \ldots, \mathbf{c}_{(j,q_n)}\}, \tag{3}$$

where $q \in \mathcal{Q}$ is one of the attributes of $\mathcal{Q}$, $n = |\mathcal{Q}|$, and $\mathbf{c}_{(j,q_i)}$ is a vector expressed as

$$\mathbf{c}_{(j,q_i)} = \left(c_j^{p_1}, c_j^{p_2}, \ldots, c_j^{p_m}\right), \tag{4}$$

where $p \in \mathcal{P}_q$, $m = \max_{q \in \mathcal{Q}} |\mathcal{P}_q|$. To identify the demographic value that has the highest preference p' for items, we introduce a function $M(\cdot)$ defined as follows

$$p'_i = M(\mathbf{c}_{(j,q_i)}) = \max_{p \in \mathcal{P}_{q_i}} c_j^p. \tag{5}$$

With this approach, we are able to construct the IUG for item j, and the member of this IUG u'_j can be represented using Remark 2 as follows

$$u'_j = M(C_j) = \{p'_1, p'_2, \ldots, p'_n\}, \tag{6}$$

where n is the length of set $\mathcal{Q}$.

3.3. User/item modeling

To address the issue of sparsity in network data, it is possible to embed all objects into a shared low-dimensional space while preserving their neighborhood proximity (Tang et al., 2015). In light of this, we opt to use an embedding technique to encode users and items, replacing the conventional one-hot encoding. We represent user i and item j as $\mathbf{p}_i$ and $\mathbf{q}_j$, respectively.

Each user's demographic information is unique, and can be represented as a set of attributes $u_i = \{p_1, p_2, \ldots, p_n\}$, similar to (6). To capture the features of users, we employ embedding techniques to encode all demographic values. We use an embedding vector $\mathbf{d}^p$ to represent each demographic value $p \in \mathcal{P}$.

[1] https://www.imdb.com/

3.4. Calculate preference score

To model the latent structures between the current user and the IUG based on demographic information, we employ the element-wise product and apply the sigmoid activation function. The result of the element-wise product is then accumulated as follows:

$$\mathbf{m}_{ij} = \sum_{g=1}^{|Q|} \sigma(\mathbf{d}^{p_g} \odot \mathbf{d}^{p'_g}), \tag{7}$$

where $\odot$ represents the element-wise product operation, and σ denotes the sigmoid function. The embedding vector $\mathbf{d}^p$ corresponds to the demographic value p, where $p_g \in u_i$ and $p'_g \in u'_j$.

To combine the embeddings of users and items, it is often useful to use connections, such as in vector concatenation (Zhang, Yang, Luan, Yang, & Chua, 2014). However, simple concatenation may not capture all the latent features between users and items (He et al., 2017). To better capture implicit features and explore the relationship between users and items, we employ a standard MLP to learn a non-linear interaction function. Thus, we can calculate the preference score between user i and item j as follows:

$$\hat{y}_{ij} = MLP(\mathbf{p}_j \oplus \mathbf{q}_i \oplus \mathbf{m}_{ij}), \tag{8}$$

where $\oplus$ denotes the vector concatenation, the $\mathbf{p}_j$ is the embedding of item j, and the $\mathbf{q}_i$ is the embedding of user i.

We use the Rectifier (ReLU) function as the activation function in the hidden layers of the MLP. At the output layer, we apply the sigmoid function to map the preference scores to the range of [0,1]. This not only provides the probability of the user selecting an item but also makes it easier to apply to the loss function.

We aim to give higher weight to users who closely match the IUG. To achieve this, we first calculate the similarity between users and the IUG using the Manhattan distance. Then, we aggregate the similarity scores s_{ij} from each demographic attribute dimension (gender, age, and occupation) to obtain an overall similarity measure:

$$s_{ij} = \sum_{q \in Q} \sum_{k=1}^{m} |\mathbf{d}_k^{p_{iq}} - \mathbf{d}_k^{p_{jq}}|, \tag{9}$$

where n represents the number of demographic attributes, and m denotes the dimension of demographics embedding. This measure of similarity captures the degree of alignment between the demographic attributes of the user and the IUG. At last, we project the similarity s to an appropriate interval and use it to adjust the preference score as follows:

$$\widetilde{y}_{ij} = \frac{\hat{y}_{ij}}{\exp(s_{ij})}. \tag{10}$$

This results in a higher preference score for users who better match the IUG of the item.

3.5. Model learning

To focus on the binary properties of the implicit data, a probabilistic approach is adopted for learning the model (He et al., 2017). In this approach, highly rated items by users are referred to as "positive" items, while items not rated by users are referred to as "negative" items. The loss function used for training the model is given by:

$$L = -\sum_{(i,j) \in \mathcal{W}} \log \widetilde{y}_{ij} - \sum_{(i,k) \in \mathcal{W}^-} \log(1 - \widetilde{y}_{ik}), \tag{11}$$

where i, j and k represent user i, a positive item j, and a negative item k, respectively. Here, $\mathcal{W}$ and $\mathcal{W}^-$ represent the sets of positive and negative samples, respectively, defined as:

$$\mathcal{W} = \{(i,j) | i \in \mathcal{I} \wedge j \in \mathcal{J}_i\}, \tag{12}$$

$$\mathcal{W}^- = \{(i,k) | i \in \mathcal{I} \wedge k \in \mathcal{J} \setminus \mathcal{J}_i\}, \tag{13}$$

where $\mathcal{J}_i$ is the set of items rated by user i, and $\mathcal{J}$ is the set of all items. The loss function is minimized during the training process to improve the model's ability to accurately predict the preference score for user–item pairs. To further encourage the embeddings of highly-rated users and the IUG of item j to be closer, we add a special regularization term to the loss function. We use the difference between the embeddings of the current user and the IUG of item j as the regularization term:

$$w_{ij} = \sum_{g=1}^{n} (\mathbf{d}^{p_{ig}} - \mathbf{d}^{p_{jg}}). \tag{14}$$

By combining (11) and (14), we obtain a new loss function:

$$\begin{aligned} L = &- \sum_{(i,j) \in \mathcal{W}} (\log \widetilde{y}_{ij} - \lambda ||w_{ij}||_1) \\ &- \sum_{(i,k) \in \mathcal{W}^-} (\log(1 - \widetilde{y}_{ik}) + \lambda ||w_{ik}||_1) \end{aligned} \tag{15}$$

where λ is the regularization coefficient.

Algorithm 1 presents the procedure of our algorithm.

Algorithm 1 Learning algorithm for IUG-CF

Input: $\mathcal{R}$: user-item interaction set; $\mathcal{D}$: demographics of all users; η: learning rate; λ: regularization coefficient; γ: expansion coefficient; Epochs: training iterations;
Output: Θ: all parameters in the model;

```
Initialize all parameters in Θ;
for j ∈ J do                                  ▷ Generate the ideal user
    for q ∈ Q do
        for p ∈ P_q do
            c_j^p ← use (1) with R, D, and γ;
        end for
    end for
    Generate the IUG via (3)–(6);
end for
U' ← all IUs;
for epoch in range(Epochs): do
    ỹ_ij ← use (7) - (10) with U' and D;
    w_ij ← use (14) with U' and D;
    L ← use (15) with ỹ_ij, w_ij, and λ;
    Use Adam with learning rate η to optimize the parameters;
end for
return all parameters in Θ;
```

4. Experiments

In this section, we present the experimental results of the IUG-CF model on three datasets. We provide details on the datasets, evaluation strategy, and parameters used in our experiments, respectively, in Sections 4.1, 4.2, and 4.3. We also describe two IUG characteristics in Section 4.4. Section 4.5 presents the internal comparison of the IUG-CF model and the comparison with existing state-of-the-art methods. Finally, in Section 4.6, we discuss and analyze the experimental results.

4.1. Datasets

In this work, three real-world datasets of different scales have been used: Movielans-100k, Movielens-1M,[2] and KuaiRec.[3] Table 2 provides a summary of the statistics of these datasets. Movielens dataset is The Movielens dataset is widely used in research on DRSs. It provides valuable data such as users' explicit ratings for movies, as well as additional user information such as gender, age, occupation, and more (Harper & Konstan, 2015). KuaiRec is a real-world dataset collected from the

[2] https://grouplens.org/datasets/movielens/
[3] https://kuairec.com/

Table 2
Dataset statistics. $\overline{|\mathcal{I}_j|}$ and $\overline{|\mathcal{J}_i|}$ represent the average actions of users and the average actions of items, respectively.

Dataset	#users	#items	#interactions	Sparsity	$\overline{\lvert\mathcal{I}_j\rvert}$	$\overline{\lvert\mathcal{J}_i\rvert}$	Demographic attributes
Movielens-100k	943	1682	100,000	93.70%	106.04	59.45	Age, Gender, and Occ.
Movielens-1M	6040	3706	1,000,209	95.53%	165.60	269.89	Age, Gender, and Occ.
KuaiRec	7176	10728	12,530,806	83.70%	1746.20	1168.05	Desensitized information

recommendation logs of the video-sharing mobile app Kuaishou (Gao et al., 2022).

In Movielens datasets, we select demographic information such as the gender, age, and occupation of the users. To ensure consistency, we converted the ages of users in the Movielens-100k dataset into the same segments as those in the Movielens-1M dataset, since the ages of users in the Movielens-1M dataset are displayed in segments. In KuaiRec, the user's demographic information is desensitized and labeled as "onehot_feat". In this paper, we chose to select the first three "onehot_feat" because their value ranges are similar to those of gender, age, and occupation in the Movielens dataset.

4.2. Evaluation metrics

We use the widely-used leave-one-out method to evaluate the performance of the recommendation algorithms (Bai, Wen, Zhang, & Zhao, 2017; Zangerle & Bauer, 2022; Zhang et al., 2018; Zhao et al., 2021). Specifically, for each user, we randomly select one item as the test item and use the remaining interactions as training data. To ensure the significance of predicting high-rating interactions (Steck, 2010), it is required that the user's rating for this test item must be above 4 (1–5). For the KuaiRec dataset, due to lack of explicit data, we set interactions with "watch_ratio $\geq$ 3" as candidates for the test item. Then, we randomly sample 100 items that are not in the user's interacted item set. We use the proposed model to rank these 101 items (the test item and 100 random items) according to the predicted probabilities. To ensure a fair comparison among all models, we have also adopted the same evaluation strategy for all baseline methods.

To evaluate the recommendation results of different algorithms, we use the Hit Ratio (HR) and Normalized Discounted Cumulative Gain (NDCG) metrics. HR@k represents the percentage of test items that appear in the top-k positions of the recommendation list out of all test items. NDCG@k is more concerned with the position, giving more weight to test items with higher positions. The formulas for calculating HR and NDCG are defined as follows, respectively:

$$\mathrm{HR}@k = \frac{\#hits@k}{\#users}, \tag{16}$$

$$\mathrm{NDCG}@k = \frac{1}{\#users}\sum_{l=1}^{\#users}\frac{1}{\log_2(p_{l,k}+1)}, \tag{17}$$

where $\#hits@k$ denotes the number of users for whom the test item appears in the top-k positions in the recommendation list, and $p_{l,k}$ is the position of the test item in the recommendation list of the lth user.

4.3. Implementation details

The IUG-CF model is implemented in Python[4] version 3.7 using Keras[5] and Tensorflow[6] version 2.5.0. The experimentation was conducted on Windows 11 operating system with Intel(R) Core(TM) i7-11800 CPU 2.30 GHz processors. The parameters of the proposed model were initialized at random using a normal distribution with a mean of 0 and a standard deviation of 0.01. We train our model using the pair-wise loss (11) and the mini-batch Adam optimizer (Kingma & Ba, 2014) with a learning rate of 0.001 and a batch size of 256. The embedding size of the user/item and demographic features is set to 512 and 128, respectively. We use a single MLP layer with a dimension of 512. The l_1-norm regularization coefficient λ is set to 0.001. For the Movielens-1M, Movielens-100k, and KuaiRec datasets, the expansion coefficient γ is set to 0.3, 0.8, and 0.5, respectively.

4.4. Verification of characteristics

In this section, we examine the two characteristics of IUG, namely "built-up" and "dynamic".

For the built-up characteristic, we compared three different schemes: the proposed scheme, the average random distribution scheme and the IUG construction scheme that considers all of the user's attributes individually (without combination). Fig. 5 shows the results of the comparison. It is clear that the proposed method reduces the diversity of IUGs compared to the method that integrates the demographics as a whole. Moreover, the IUG constructed without combination is nearly equally divided based on the distribution of each attribute, while the method with combination enlarges the distribution difference of each attribute. It is important to note that the proposed method excludes ambiguous factors like "other" in the occupation distribution, resulting in more accurate identification of the best users for an item.

To evaluate the impact of the dynamic characteristic on recommendation performance, we conducted a study comparing several schemes under different training sample percentages. Firstly, we sorted all historical rating records by time and used the earliest portion in proportion to construct IUGs. Fig. 6 illustrates how the number of IUG categories varies with the sampling percentage. It is evident that the category of IUGs built using the combination method becomes less and undergoes more significant changes as the number of samples increases. Fig. 7 shows the recommendation performance of the three methods at different sampling percentages. It is observed that as the learning samples increase, the IUG constructed by our model becomes more accurate. The method of constructing IUGs with combinations displays better recommendation performance, indicating that the proposed method can eliminate inaccurate IUGs of items. In other words, the proposed method can more accurately identify the most suitable user group for an item.

4.5. Comparison with baselines

In Section 4.5.1, we present five versions of the IUG-CF model that we used in internal experiments to explore the optimal model structure. We compare the performance of these models with six state-of-the-art methods, which are described in Section 4.5.2.

4.5.1. Our models

We conducted an internal comparison by replacing or removing parts of the model's structure to contrast its performance. The internal comparison included the following models:

- IUG-CF: The proposed model.
- IUG-CF-W: This model treats demographic information as an inseparable unit in the process of building IUGs.
- IUG-CF-S: This model uses only partial demographic information (age in this case) in the process of building IUGs.
- IUG-CF-A: This model removes the adjustment of predicted scores based on the similarity between the IUG and users (Eqs. (9)–(10)).

[4] https://www.python.org/
[5] https://keras.io/
[6] https://www.tensorflow.org/

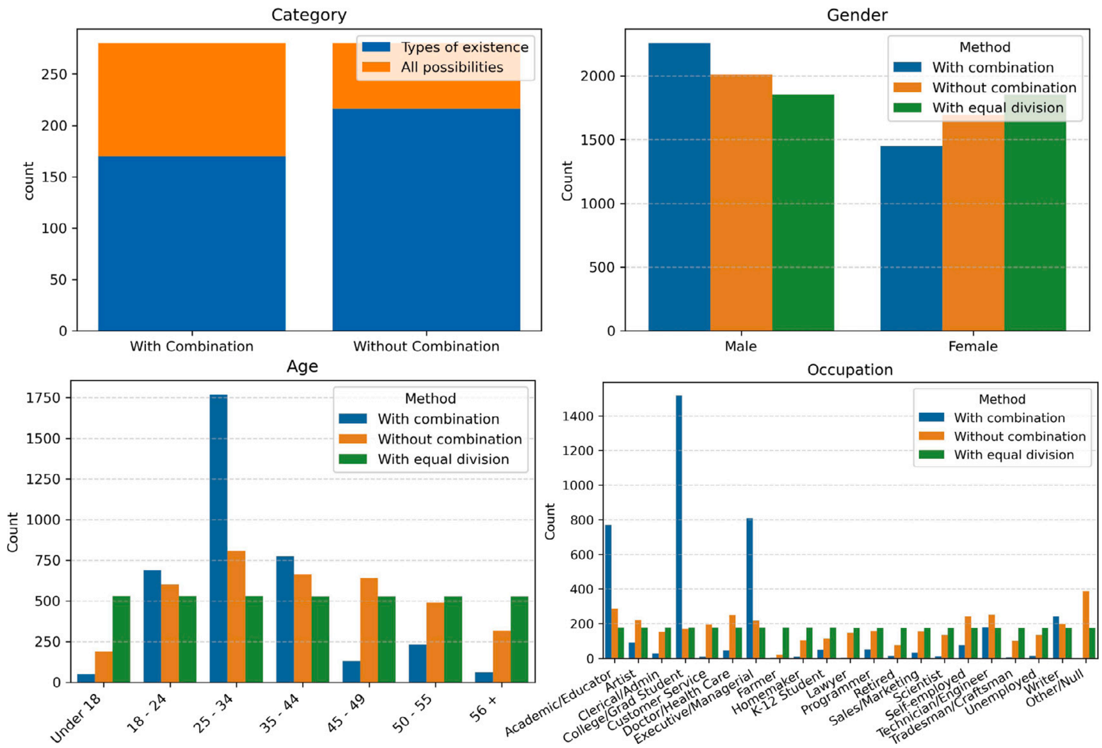


Fig. 5. Comparison of distribution of each attribute in Movielens-1M in different construction methods.

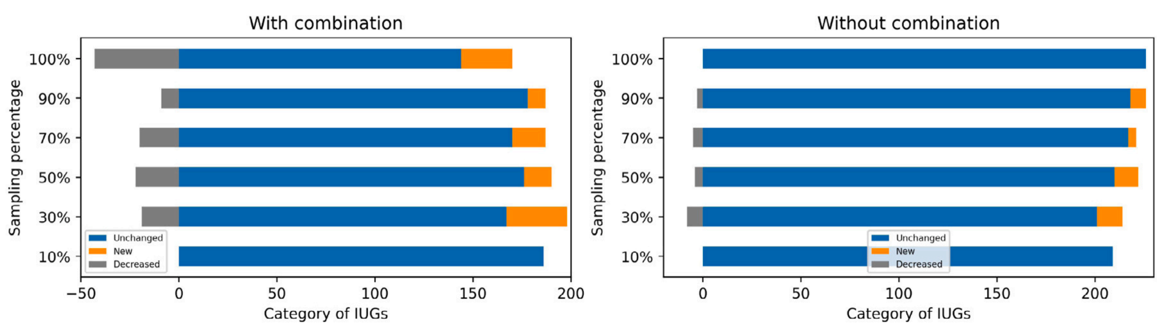


Fig. 6. Variation chart of IUG categories with sampling percentage.

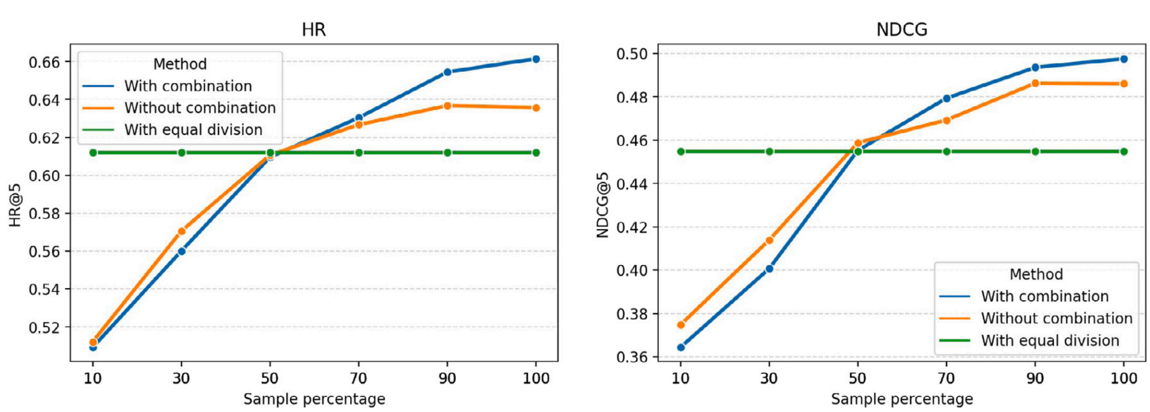


Fig. 7. Variation trend of evaluation metrics of three methods in Movielens-1M with different sampling percentages.

- IUG-CF-L: This model removes the modification of the loss function based on the distance between the IUG and the user (Eqs. (14)–(15)).
- IUG-CF-M: This model incorporates the element-wise product $\mathbf{n}$ of user and item embeddings, and modifies Eq. (8) to $\hat{y}_{ij} = MLP(\mathbf{p}_j \oplus \mathbf{q}_i \oplus \mathbf{m}_{ij} \oplus \mathbf{n_{ij}})$.
- IUG-CF-B: This model set a_j^q to 0 in Eq. (1).

4.5.2. Baseline models

We mainly compare our model with demographic-based models and representative DL-based models:

- ItemAve (Safoury & Salah, 2013): This is a demographic-based model designed to address the cold-start problem. It computes the average rating of items for each demographic attribute and then recommends items with the highest rating on the corresponding attribute, based on the user's demographics.
- NeuMF (He et al., 2017): This is a CF model that combines MF and MLP to model both linear and nonlinear user–item interactions. It is capable of learning arbitrary functions from data.
- LRML (Tay et al., 2018): This is a DL-based model with implicit feedback that introduces a latent relational metric learning approach that learns exclusive and optimal relational translations for user–item interactions.
- Wide&Deep (Cheng et al., 2016): This is a model proposed by Google that maps features such as demographics into embeddings and simultaneously trains a linear model and a deep model to improve the generalization ability of the model.
- CFFNN (Yu et al., 2021): This is a DL-based model that combines user and item features to comprehend user preference for items.
- HybridNNMF (Zeng et al., 2021): This is a hybrid model that combines a global neural network with several local neural blocks to capture both linear and nonlinear relationships between users and items.
- IFDNAMF (Tian et al., 2021): This is a DL-based recommendation model, which introduces attribute information during information fusion and uses deep neural networks to learn high-order interactions between users and items.
- NHF (Sarker & Matin, 2021): This is a hybrid model that combines MF and MLP and incorporates user demographics and item auxiliary information.
- KTUP (Cao, Wang, He, Hu, & Chua, 2019): This model concurrently learns both recommendation and knowledge graph completion, with a focus on transferring relation information to enhance the understanding of user preferences.
- CoupledCF (Zhang et al., 2018): This is a model that jointly learns explicit and implicit coupling within and between users and items. It combines user demographics with item extra information, using a convolutional neural network to learn user–item coupling.

4.6. Results and discussion

As mentioned before, the IUG-CF model has five customized versions. IUG-CF-W and IUG-CF-S differ in their approach to building the IUGs, while IUG-CF-A and IUG-CF-L exclude the steps of prediction score smoothing and loss function adjustment, respectively. Fig. 8 illustrates that IUG-CF performs better than the other four versions on the Movielens-1M dataset. This suggests that combining demographic information through split-combining yields better user–item matching than split-only (IUG-CF-S) or no-split (IUG-CF-W) methods. Comparing IUG-CF with IUG-CF-A and IUG-CF-L shows that our approach of smoothing prediction scores and adjusting the loss function is effective. Furthermore, comparing IUG-CF with IUG-CF-M, the results show that modeling more linear interactions with the GMF (He et al., 2017) structure does not improve the recommendation performance. One possible reason is that in the IUG-CF model, the inner product between the embeddings of demographic information has already captured sufficient linear interactions. Finally, the results of the IUG-CF-B model demonstrate that utilizing Bayesian averaging indeed creates a fair comparison environment and leads to the construction of more accurate IUGs mapping for items.

Further, the proposed model is compared with existing methods listed in Section 4.5.2. Table 3 provides a comparison of all methods for generating a top-k recommendation list, which the best-performing marked in bold. It is evident that IUG-CF outperforms all other methods on both datasets in terms of two different top-k values. Compared to NeuMF, the NDCG@10 indicator increased by 22.79%, indicating that incorporating demographic information into the neural network is effective.

Generally, models that combine demographic information with deep learning embed demographics and concatenate these embeddings with user or item embeddings. However, using user demographics alone as feature input limits the value of demographic information. Models like Wide&Deep, IFDNACF, NHF, and CoupledCF map user demographics as embeddings and integrate them into a DL-based recommender model. The results confirm the rationale behind constructing IUGs and learning the user-IUG connections for demographic recommendation. They also suggest that data mining with demographic information before combining it with deep learning improves recommendation performance.

To evaluate the effectiveness of the proposed model across various domains and datasets of varying sizes, we conducted experiments using the KuaiRec dataset. KuaiRec is a large-scale and denser short video dataset. For comparison, we selected the NeuMF and CoupledCF models and refined the evaluation granularity to observe performance changes within the range of K from 2 to 10.

We specifically chose these two models for the following reasons: NeuMF is a representation of the most classical baseline model in neural collaborative filtering, while CoupledCF closely mirrors the performance of the proposed model among the baseline models. Table 4 displays the results of the comparison. The results indicate that as K decreases, the improvement in both indicators compared to NeuMF becomes increasingly significant. This suggests that leveraging demographic information is effective in generating more personalized and accurate recommendations. Therefore, for video recommendation, it is also effective to map items with demographic information to construct IUGs. In summary, the results validate the effectiveness of the proposed IUG-CF model for demographic recommendation.

5. Conclusion and future work

This study proposes the concept of IUG, which is a dynamic feature that describes the user group for which an item is most suitable for recommendation. To validate the effectiveness and rationality of IUG, we propose an IUG-based neural CF model called IUG-CF and conduct experiments on three real-world datasets. The experiments first confirm IUG's "built-up" and "dynamic" characteristics. Secondly, internal comparisons are deployed to determine the optimal structure for IUG-CF. Finally, the IUG-CF model is compared to existing state-of-the-art models, and the experimental results demonstrate that combining a DL-based recommendation model with the IUG is effective.

From the viewpoint of methodological, the following research topics deserve further investigation:

1. Latent factor analysis. Latent factor analysis is a highly effective method for analyzing high-dimensional sparse data in recommender systems. It is commonly applied in deep learning-based recommender systems to model user–item interactions and capture latent features that represent user preferences and item characteristics (Sarker & Matin, 2021; Wang, Hong, &

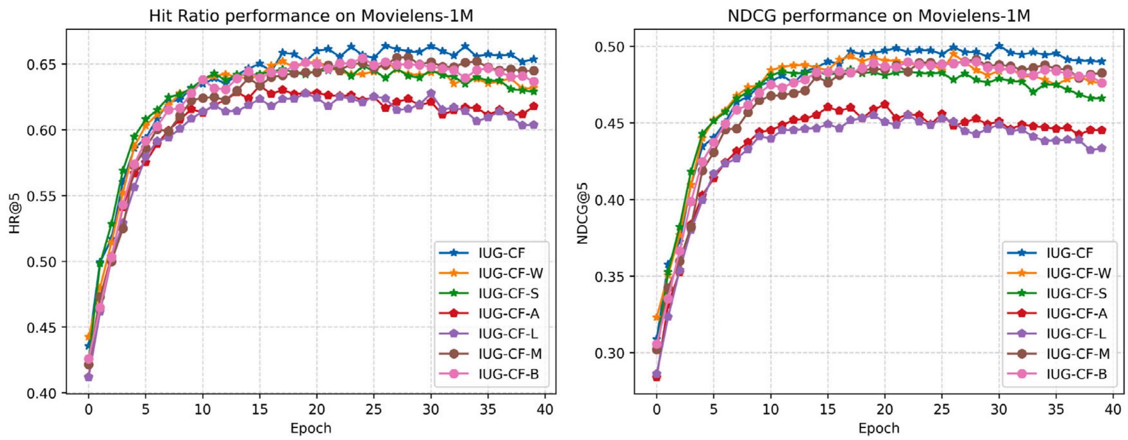


Fig. 8. The internal comparison on the MovieLens-1M dataset using HR@5 and NDCG@5 metrics.

Table 3
Performance comparison of different models on Movielens-100k and Movielens-1M. The standard error is shown in the brackets and bold values indicate the best results.

Datasets	Movielens-100k		Movielens-1M	
Metrics	HR@10	NDCG@10	HR@10	NDCG@10
ItemAve	0.3278	0.1632	0.4822	0.1820
NeuMF	0.6807	0.4201	0.7260	0.4415
LRML	0.6363	0.3665	0.7281	0.4303
HybridNNMF	0.7260	0.4200	0.7380	0.4560
Wide&Deep	0.7400	0.4509	0.7387	0.4490
CFFNN	0.7012	0.4077	0.7511	0.4582
IFDNAMF	0.7481	0.4591	0.7558	0.4680
NHF	0.7784	0.4915	0.7715	0.4950
KTUP	0.7614	0.5138	0.7818	0.5156
CoupledCF	0.7895	0.5117	0.8003	0.5388
IUG-CF	**0.7931** (.0017)	**0.5399** (.0040)	**0.8022** (.0013)	**0.5408** (.0017)

Table 4
Performance Comparison: Proposed Model, NeuMF, and CoupledCF on KuaiRec. The standard error is shown in the brackets.

K	HR@K			NDCG@K		
	NeuMF	CoupledCF	IUG-CF	NeuMF	CoupledCF	IUG-CF
2	0.3948 (.0042)	0.4612 (.0019)	**0.4809** (.0027)	0.3487 (.0017)	0.4174 (.0017)	**0.4273** (.0005)
4	0.5578 (.0014)	0.6288 (.0031)	**0.6391** (.0009)	0.4235 (.0035)	0.4931 (.0015)	**0.5027** (.0004)
6	0.6650 (.0018)	0.7265 (.0019)	**0.7310** (.0015)	0.4635 (.0012)	0.5238 (.0013)	**0.5374** (.0005)
8	0.7351 (.0015)	0.7922 (.0022)	**0.7943** (.0019)	0.4851 (.0019)	0.5452 (.0010)	**0.5580** (.0006)
10	0.7918 (.0023)	0.8405 (.0011)	**0.8418** (.0009)	0.5015 (.0020)	0.5602 (.0010)	**0.5713** (.0006)

Hong, 2022; Zeng et al., 2021). Therefore, integrating it into the IUG-based model is likely to enhance the recommendation performance.

2. High-level feature extraction. The user demographic information used in our work comes from the dataset. Deep neural network can extract higher-level user features to replace user demographic information to deal with incomplete demographics in the dataset (Bello, Nápoles, Sánchez, Bello, & Vanhoof, 2020; Li, Jiang, Liu, Wang, & Li, 2022).
3. Popularity and selection bias mitigation. Popularity bias arises when certain user groups focus more on popular items, while selection bias occurs when explicit rating influences the IUG construction. Mitigating popularity bias and selection bias is crucial for constructing an effective IUG-based model (Chen, Dong, et al., 2023; Shi, Li, Ding, & Bu, 2023).
4. Sequential recommendation. The integration of sequential recommendation and demographic information merits further investigation (Noorian, 2024). User preferences evolve over time, influenced by their historical interactions. By introducing interaction sequences, additional features in the changing process of IUG can be mined to offer more accurate recommendations.

CRediT authorship contribution statement

Zi-Feng Peng: Conceptualization, Methodology, Software, Data curation, Visualization, Writing – original draft. **Heng-Ru Zhang:** Conceptualization, Methodology, Validation, Writing – review & editing. **Fan Min:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the link to my code at the manuscript.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61902328) and the Applied Basic Research Project of

Science and Technology Bureau of Nanchong City, China (SXHZ040 and SXHZ051).

Code availability

The source code used to perform all the experiments is publicly available. The link is the following: https://github.com/zhanghrswpu/IUG-CF.

References

Al-Shamri, M. Y. H. (2016). User profiling approaches for demographic recommender systems. *Knowledge-Based Systems*, *100*, 175–187. http://dx.doi.org/10.1016/j.knosys.2016.03.006.

Al-Shamri, M. Y. H., & Bharadwaj, K. K. (2007). A compact user model for hybrid movie recommender system. In *ICCIMA, vol. 1* (pp. 519–524). http://dx.doi.org/10.1109/ICCIMA.2007.15.

Bai, T., Wen, J.-R., Zhang, J., & Zhao, W. X. (2017). A neural collaborative filtering model with interaction-based neighborhood. In *CIKM* (pp. 1979–1982). http://dx.doi.org/10.1145/3132847.3133083.

Bello, M., Nápoles, G., Sánchez, R., Bello, R., & Vanhoof, K. (2020). Deep neural network to extract high-level features and labels in multi-label classification problems. *Neurocomputing*, *413*, 259–270. http://dx.doi.org/10.1016/j.neucom.2020.06.117.

Bobadilla, J., González-Prieto, Á., Ortega, F., & Lara-Cabrera, R. (2021). Deep learning feature selection to unhide demographic recommender systems factors. *Neural Computing and Applications*, *33*, http://dx.doi.org/10.1007/s00521-020-05494-2.

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, *46*, 109–132. http://dx.doi.org/10.1016/j.knosys.2013.03.012.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and User-adapted Interaction*, *12*(4), 331–370. http://dx.doi.org/10.1023/A:1021240730564.

Cao, Y., Wang, X., He, X., Hu, Z., & Chua, T.-S. (2019). Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences. In *WWW* (pp. 151–161). http://dx.doi.org/10.1145/3308558.3313705.

Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, X. (2023). Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, *41*(3), 1–39. http://dx.doi.org/10.1145/3564284.

Chen, L., Wu, L., Zhang, K., Hong, R., Lian, D., Zhang, Z., et al. (2023). Improving recommendation fairness via data augmentation. In *WWW* (pp. 1012–1020). http://dx.doi.org/10.1145/3543507.3583341.

Cheng, H.-T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., et al. (2016). Wide & deep learning for recommender systems. In *DLRS* (pp. 7–10). http://dx.doi.org/10.1145/2988450.2988454.

Covington, P., Adams, J., & Sargin, E. (2016). Deep neural networks for youtube recommendations. In *RecSys* (pp. 191–198). http://dx.doi.org/10.1145/2959100.2959190.

Gao, C., Li, S., Lei, W., Chen, J., Li, B., Jiang, P., et al. (2022). KuaiRec: A fully-observed dataset and insights for evaluating recommender systems. In *CIKM* (pp. 540–550). http://dx.doi.org/10.1145/3511808.3557220.

Harper, F. M., & Konstan, J. A. (2015). The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, *5*(4), 1–19. http://dx.doi.org/10.1145/2827872.

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T.-S. (2017). Neural collaborative filtering. In *WWW* (pp. 173–182). http://dx.doi.org/10.1145/3038912.3052569.

Jameson, A. (2004). More than the sum of its members: challenges for group recommender systems. In *AVI* (pp. 48–54). http://dx.doi.org/10.1145/989863.989869.

Khojamli, H., & Razmara, J. (2021). Survey of similarity functions on neighborhood-based collaborative filtering. *Expert Systems with Applications*, *185*, http://dx.doi.org/10.1016/j.eswa.2021.115482.

Kim, D., Park, C., Oh, J., Lee, S., & Yu, H. (2016). Convolutional matrix factorization for document context-aware recommendation. In *RecSys* (pp. 233–240). http://dx.doi.org/10.1145/2959100.2959165.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. http://dx.doi.org/10.48550/arXiv.1412.6980, arXiv preprint arXiv:1412.6980.

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. (1997). Grouplens: Applying collaborative filtering to usenet news. *Communications of the ACM*, *40*(3), 77–87. http://dx.doi.org/10.1145/245108.2451264.

Krulwich, B. (1997). Lifestyle finder: Intelligent user profiling using large-scale demographic data. *AI Magazine*, *18*(2), 37.

Lekakos, G., & Giaglis, G. M. (2007). A hybrid approach for improving predictive accuracy of collaborative filtering algorithms. *User Modeling and User-Adapted Interaction*, *17*(1), 5–40. http://dx.doi.org/10.1007/s11257-006-9019-0.

Li, X., Jiang, H., Liu, Y., Wang, T., & Li, Z. (2022). An integrated deep multiscale feature fusion network for aeroengine remaining useful life prediction with multisensor data. *Knowledge-Based Systems*, *235*, Article 107652. http://dx.doi.org/10.1016/j.knosys.2021.107652.

Liu, Y., Qu, H., Chen, W., & Mahmud, S. H. (2019). An efficient deep learning model to infer user demographic information from ratings. *IEEE Access*, *7*, 53125–53135. http://dx.doi.org/10.1109/ACCESS.2019.2911720.

Masthoff, J. (2010). Group recommender systems: Combining individual models. In *Recommender systems handbook* (pp. 677–702). http://dx.doi.org/10.1007/978-0-387-85820-3_21.

Masthoff, J. (2015). Group recommender systems: aggregation, satisfaction and group attributes. In *Recommender systems handbook* (pp. 743–776). http://dx.doi.org/10.1007/978-1-4899-7637-6_22.

Mooney, R. J., & Roy, L. (2000). Content-based book recommending using learning for text categorization. In *DL* (pp. 195–204). http://dx.doi.org/10.1145/336597.336662.

Noorian, A. (2024). A BERT-based sequential POI recommender system in social media. *Computer Standards & Interfaces*, *87*, Article 103766. http://dx.doi.org/10.1016/j.csi.2023.103766.

Pandey, A. K., & Rajpoot, D. S. (2016). Resolving cold start problem in recommendation system using demographic approach. In *ICSC* (pp. 213–218). http://dx.doi.org/10.1109/ICSPCom.2016.7980578.

Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, *13*(5), 393–408. http://dx.doi.org/10.1023/A:1006544522159.

Pereira, A. L. V., & Hruschka, E. R. (2015). Simultaneous co-clustering and learning to address the cold start problem in recommender systems. *Knowledge-Based Systems*, *82*, 11–19. http://dx.doi.org/10.1016/j.knosys.2015.02.016.

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, *92*, http://dx.doi.org/10.1080/01621459.1997.10473615.

Renjith, S., Sreekumar, A., & Jathavedan, M. (2020). An extensive study on the evolution of context-aware personalized travel recommender systems. *Information Processing & Management*, *57*(1), Article 102078. http://dx.doi.org/10.1016/j.ipm.2019.102078.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). Grouplens: An open architecture for collaborative filtering of netnews. In *CSCW* (pp. 175–186). http://dx.doi.org/10.1145/192844.192905.

Rich, E. (1979). User modeling via stereotypes. *Cognitive Science*, *3*(4), 329–354. http://dx.doi.org/10.1207/s15516709cog0304_3.

Sacharidis, D. (2019). Top-n group recommendations with fairness. In *SAC* (pp. 1663–1670). http://dx.doi.org/10.1145/3297280.3297442.

Safoury, L., & Salah, A. (2013). Exploiting user demographic attributes for solving cold-start problem in recommender system. *Lecture Notes on Software Engineering*, *1*(3), 303–307. http://dx.doi.org/10.7763/LNSE.2013.V1.66.

Sahu, A. K., & Dwivedi, P. (2019). User profile as a bridge in cross-domain recommender systems for sparsity reduction. *Applied Intelligence*, *49*(7), 2461–2481. http://dx.doi.org/10.1007/s10489-018-01402-3.

Sarker, M. R. I., & Matin, A. (2021). A hybrid collaborative recommendation system based on matrix factorization and deep neural network. In *ICICT4SD* (pp. 371–374). http://dx.doi.org/10.1109/ICICT4SD50815.2021.9397027.

Shen, R.-P., Zhang, H.-R., Yu, H., & Min, F. (2019). Sentiment based matrix factorization with reliability for recommendation. *Expert Systems with Applications*, *135*, 249–258. http://dx.doi.org/10.1016/j.eswa.2019.06.001.

Shi, L., Li, S., Ding, X., & Bu, Z. (2023). Selection bias mitigation in recommender system using uninteresting items based on temporal visibility. *Expert Systems with Applications*, *213*, Article 118932. http://dx.doi.org/10.1016/j.eswa.2022.118932.

Souali, K., El Afia, A., & Faizi, R. (2011). An automatic ethical-based recommender system for e-commerce. In *ICMCS* (pp. 1–4). http://dx.doi.org/10.1109/ICMCS.2011.5945631.

Steck, H. (2010). Training and testing of recommender systems on data missing not at random. In *SIGKDD* (pp. 713–722). http://dx.doi.org/10.1145/1835804.1835895.

Tahmasebi, F., Meghdadi, M., Ahmadian, S., & Valiallahi, K. (2021). A hybrid recommendation system based on profile expansion technique to alleviate cold start problem. *Multimedia Tools and Applications*, *80*, 2339–2354. http://dx.doi.org/10.1007/s11042-020-09768-8.

Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., & Mei, Q. (2015). Line: Large-scale information network embedding. In *WWW* (pp. 1067–1077). http://dx.doi.org/10.1145/2736277.2741093.

Tay, Y., Anh Tuan, L., & Hui, S. C. (2018). Latent relational metric learning via memory-based attention for collaborative ranking. In *WWW* (pp. 729–739). http://dx.doi.org/10.1145/3178876.3186154.

Tian, Z., Pan, L., Yin, P., & Wang, R. (2021). Information fusion-based deep neural attentive matrix factorization recommendation. *Algorithms*, *14*(10), 281. http://dx.doi.org/10.3390/a14100281.

Tiwari, V., Ashpilaya, A., Vedita, P., Daripa, U., & Paltani, P. P. (2020). Exploring demographics and personality traits in recommendation system to address cold start problem. In *ICT4SD* (pp. 361–369). http://dx.doi.org/10.1007/978-981-15-0936-0_37.

Van Meteren, R., & Van Someren, M. (2000). Using content-based filtering for recommendation. In *Proceedings of the machine learning in the new information age: MLnet/ECML2000 workshop, vol. 30* (pp. 47–56).

Vozalis, M. G., & Margaritis, K. G. (2007). Using SVD and demographic data for the enhancement of generalized collaborative filtering. *Information Sciences*, *177*(15), 3017–3037. http://dx.doi.org/10.1016/j.ins.2007.02.036.

Wang, H., Hong, Z., & Hong, M. (2022). Research on product recommendation based on matrix factorization models fusing user reviews. *Applied Soft Computing*, *123*, Article 108971. http://dx.doi.org/10.1016/j.asoc.2022.108971.

Wang, D., Liang, Y., Xu, D., Feng, X., & Guan, R. (2018). A content-based recommender system for computer science publications. *Knowledge-Based Systems*, *157*, 1–9. http://dx.doi.org/10.1016/j.knosys.2018.05.001.

Xue, F., He, X., Wang, X., Xu, J., Liu, K., & Hong, R. (2019). Deep item-based collaborative filtering for top-n recommendation. *ACM Transactions on Information Systems*, *37*, http://dx.doi.org/10.1145/3314578.

Yang, X., & Zhang, Z. (2013). Combining prestige and relevance ranking for personalized recommendation. In *CIKM* (pp. 1877–1880). http://dx.doi.org/10.1145/2505515.2507885.

Yassine, A., Mohamed, L., & Al Achhab, M. (2021). Intelligent recommender system based on unsupervised machine learning and demographic attributes. *Simulation Modelling Practice and Theory*, *107*, Article 102198. http://dx.doi.org/10.1016/j.simpat.2020.102198.

Yu, R., Ye, D., Wang, Z., Zhang, B., Oguti, A. M., Li, J., et al. (2021). CFFNN: Cross feature fusion neural network for collaborative filtering. *IEEE Transactions on Knowledge and Data Engineering*, *34*(10), 4650–4662. http://dx.doi.org/10.1109/TKDE.2020.3048788.

Zangerle, E., & Bauer, C. (2022). Evaluating recommender systems: survey and framework. *ACM Computing Surveys*, *55*(8), 1–38. http://dx.doi.org/10.1145/3556536.

Zeng, W., Fan, G., Sun, S., Geng, B., Wang, W., Li, J., et al. (2021). Collaborative filtering via heterogeneous neural networks. *Applied Soft Computing*, *109*, Article 107516. http://dx.doi.org/10.1016/j.asoc.2021.107516.

Zhang, Q., Cao, L., Zhu, C., Li, Z., & Sun, J. (2018). Coupledcf: Learning explicit and implicit user-item couplings in recommendation for deep collaborative filtering. In *IJCAI* (pp. 3662–3668). http://dx.doi.org/10.5555/3304222.3304277.

Zhang, H.-R., Min, F., Zhang, Z.-H., & Wang, S. (2019). Efficient collaborative filtering recommendations with multi-channel feature vectors. *International Journal of Machine Learning and Cybernetics*, *10*(5), 1165–1172. http://dx.doi.org/10.1007/s13042-018-0795-8.

Zhang, H., Yang, Y., Luan, H., Yang, S., & Chua, T.-S. (2014). Start from scratch: Towards automatically identifying, modeling, and naming visual attributes. In *MM* (pp. 187–196). http://dx.doi.org/10.1145/2647868.2654915.

Zhao, X. W., Guo, Y., He, Y., Jiang, H., Wu, Y., & Li, X. (2014). We know what you want to buy: a demographic-based system for product recommendation on microblogs. In *SIGKDD* (pp. 1935–1944). http://dx.doi.org/10.1145/2623330.2623351.

Zhao, W. X., Li, S., He, Y., Wang, L., Wen, J.-R., & Li, X. (2016). Exploring demographic information in social media for product recommendation. *Knowledge and Information Systems*, *49*, 61–89. http://dx.doi.org/10.1007/s10115-015-0897-5.

Zhao, W. X., Mu, S., Hou, Y., Lin, Z., Chen, Y., Pan, X., et al. (2021). Recbole: Towards a unified, comprehensive and efficient framework for recommendation algorithms. In *Proceedings of the 30th Acm international conference on information & knowledge management* (pp. 4653–4664). http://dx.doi.org/10.1145/3459637.3482016.