

FRC: 一种基于图嵌入对抗学习公平表示的推荐算法

李昊洋

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月30日; 录用日期: 2024年5月23日; 发布日期: 2024年5月31日

摘要

推荐系统的目标是过滤和筛选信息, 从而预测用户最感兴趣的内容。但公平性问题可能会对推荐结果产生影响。当前的公平性推荐模型大多数假设用户实体之间的独立性, 没有考虑实体之间相关的情况下该如何缓解公平性问题。此外, 在已有的少数利用图结构缓解公平性问题的研究中, 仅利用节点嵌入学习数据表示, 而用户的敏感信息可能会被局部图结构暴露。针对上述两个问题, 本文提出了一种基于图嵌入对抗学习实现公平推荐的模型FRC, 适用于任何将用户和项目嵌入作为输入的推荐任务, 该模型将用户-项目数据映射为双向图, 结合用户级和节点级的嵌入表示, 利用对抗学习消除敏感特征以获得推荐任务中的公平表示。在两个真实世界数据集上的对比实验结果表明, 本文提出的模型在推荐任务中为用户生成了更公平的推荐效果。

关键词

对抗学习, 公平表示学习, 图嵌入, 推荐系统

FRC: A Recommendation Algorithm Based on Graph Embedding Adversarial Learning for Fair Representation

Haoyang Li

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 30th, 2024; accepted: May. 23rd, 2024; published: May. 31st, 2024

Abstract

The goal of recommendation systems is to filter and select information in order to predict the content that users are most interested in. However, fairness issues may impact the recommendation results. Most existing fairness recommendation models assume independence between user entities and do not consider how to mitigate fairness issues when entities are correlated. Additionally, in a few existing studies that utilize graph structures to address fairness issues, only node embeddings are used for learning data representations, potentially exposing sensitive user information through local graph structures. To address these two issues, this paper proposes a model called FRC, based on graph embedding adversarial learning, to achieve fair recommendations. It is applicable to any recommendation task that takes user and item embeddings as input. The model maps user-item data into a bidirectional graph, combines user-level and node-level embeddings, and utilizes adversarial learning to eliminate sensitive features for fair representation in recommendation tasks. Comparative experiments on two real-world datasets demonstrate that the proposed model generates fairer recommendations for users in recommendation tasks.

Keywords

Adversarial Learning, Fair Representation Learning, Graph Embedding, Recommendation System

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在如今信息爆炸的时代，推荐系统广泛应用于各种场景，为用户提供个性化建议。近期有一些研究表明，推荐结果常受算法偏见和不公平因素影响[1] [2] [3]。例如，职业推荐时，同等条件的男女候选人，系统更偏向推荐男性[4]；在社交网络推荐中，结果常偏向多数群体，限制了少数群体的社会影响力[5]。因此，研究推荐系统中存在的不公平问题变得至关重要。

公平表示学习因其简单普适的特点受到了广泛关注，其主要任务是学习数据的表示，并利用对抗学习技术过滤掉数据表示中隐藏的敏感信息，以实现公平推荐的目标。但有大规模的研究表明，用户的敏感特征(性别、年龄)是可以从其他相似的用户行为特征中推断出来的，大多数基于公平性的推荐任务假设实体之间的独立性，并没有建模与其他实体相关情况下该如何缓解推荐不公平问题。在推荐系统中，将数据映射为图结构，可以聚合节点之间的关联信息，获得更高维度的信息表示，以解决上述问题。因此，基于图的视角学习推荐中的公平表示是至关重要的。

研究认为用户和项目的数据可以天然形成用户 - 项目双向图的结构[6] [7]，在学习用户和项目嵌入表示时引入图结构，并结合神经图网络模型，会得到优质的推荐性能[8]。此外，在已有的少数将图结构与公平性推荐结合的研究中，仅利用节点嵌入学习数据表示，独立的对每个节点的敏感特征进行过滤，用户的敏感信息可能会被其局部图结构所暴露。因此本文将为用户建立以自我为中心的图网络，聚合用户邻域信息，学习用户级的数据表示，结合节点层次的数据表示，利用图对抗学习过滤敏感特征，以缓解推荐系统中的不公平问题。

本文专注于研究公平表示学习，旨在消除表示学习中的用户敏感特征[3] [9]。公平性系统的要求为隐

藏用户的敏感特征，如性别、年龄和职业，以确保推荐结果与用户的敏感特征无关。鉴于推荐系统在实际应用中的多样性和复杂性，推荐模型应具有不可知性。故本文提出的模型适用于任何以用户和项目嵌入作为输入的推荐任务。通过定义敏感特征集，从图的视角学习过滤空间，生成屏蔽敏感特征的用户嵌入表示。具体地说，本文模型学习了每个敏感特征的过滤器，将原始用户和项目嵌入表示转换到过滤空间中，并将每个用户表示为一个自中心网络图，在对抗训练中学习数据表示，消除以用户为中心的图结构中暴露出的敏感特征；鉴别器负责预测相应的敏感特征，与过滤器共同训练，通过对抗学习使得过滤器和鉴别器二者相互学习，从而实现更公平的推荐。此外，因项目曝光偏差的存在，在数据采样过程中，计算项目的曝光度，通过负采样策略提高推荐准确性。

2. 问题定义

2.1. 推荐任务

推荐系统旨在向用户推荐其最有可能感兴趣的项目。在推荐任务中，定义 $U = \{u_1, u_2, u_3, \dots, u_m\}$ 为用户集合， $V = \{v_1, v_2, v_3, \dots, v_n\}$ 为项目集合， m 和 n 表示用户和项目的数量。用户与项目的交互历史定义为矩阵 $\mathbf{R} \in \mathbb{R}^{M \times N}$ ， r_{uv} 为用户 u 对项目 v 的评分，若用户 u 于项目 v 没有交互则 $r_{uv} = 0$ 。将用户 - 项目数据构建为双向图 $\mathcal{G} = \langle U \cup V, \mathbf{A} \rangle$ ， \mathbf{A} 是基于交互矩阵 \mathbf{R} 构建的，可以表示为：

$$\mathbf{A} = \begin{bmatrix} \mathbf{R} & \mathbf{0}^{M \times M} \\ \mathbf{0}^{M \times N} & \mathbf{R}^T \end{bmatrix} \quad (1)$$

通过编码学习用户和项目的嵌入表示 $E = \text{Enc}(\mathcal{G}) = [e_1, \dots, e_u, \dots, e_v, \dots, e_{M+N}]$ 。使用 \hat{r}_{uv} 表示对于用户 - 项目对的预测偏好得分，可以表示为：

$$\hat{r}_{uv} = e_u^T \times e_v \quad (2)$$

2.2. 公平性任务

与之前的研究一致[2] [10]，实现推荐任务公平性最直接的办法消除用户数据中的隐私敏感特征，例如年龄、职业、性别等。假设 $\mathbf{X} \in \mathbb{R}^{K \times M}$ 表示用户的敏感特征集，为了保证推荐结果不受用户敏感特征的影响，对于给定用户 u ，任意 $v \in V$ 对于用户 - 项目对预测结果的评分函数 \hat{r}_{uv} 都独立于用户敏感特征 X_u ，形式上，将公平性任务定义为：

$$I(f_u, X_u) = 0, \forall u \in U \quad (3)$$

其中 $I(\cdot)$ 表示两个变量的交互信息， f_u 表示过滤敏感特征后的用户嵌入表示。

3. 方法

本文提出的模型如图 1 所示。该模型的目标是在过滤空间中实现模型不可知的公平表示学习，也就是该模型可以应用于将用户和项目嵌入作为输入的任意推荐任务中，这里公平性的要求指的是推荐结果不受用户敏感特征集 $\mathbf{X} \in \mathbb{R}^{K \times M}$ 的影响。

本文模型使用对抗学习实现公平性，对抗网络由过滤器网络 \mathcal{F}^k 和鉴别器网络 \mathcal{D}^k 组成。本节将重点介绍过滤器结构、推荐公平性的建模以及对抗训练过程。

3.1. 过滤器结构

使用过滤器网络结构来过滤原始嵌入，其具有两个目标：1) 代表用户的个性化偏好，以保证推荐结果的准确性；2) 过滤后的嵌入表示不会泄露用户敏感特征相关的信息。

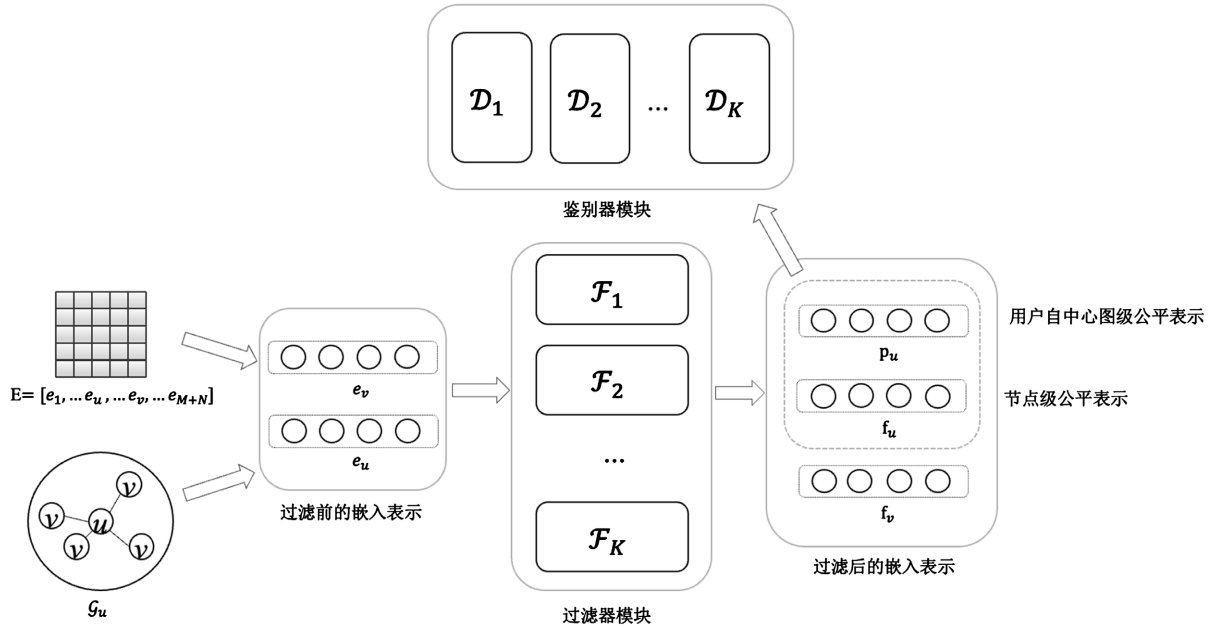


Figure 1. Overall framework of the model
图 1. 模型总体框架

给定原始嵌入矩阵 E 和敏感特征集 X ，过滤器结构包含 K 个子过滤网络以消除用户敏感特征信息，将用户节点和项目节点从原始嵌入空间 E 转换至过滤后空间 $F = \mathcal{F}(\mathcal{G}, E, X)$ 。过滤网络 \mathcal{F} 包含 K 个子过滤器，也就是 $\mathcal{F} = [\mathcal{F}^k]_{k=1}^K$ ，每个子过滤器被训练为过滤第 k 个敏感特征。过滤后的用户或项目嵌入表示如下：

$$f_i = \frac{\sum_{k=1}^K \mathcal{F}^k(e_i)}{K} \tag{4}$$

其中， e_i 表示原始嵌入矩阵中的用户或项目表示。

过滤后的嵌入表示 f_i 消除了用户敏感特征信息，预测偏好得分 \hat{r}_{uv} 此时可以被表示为：

$$\hat{r}_{uv} = f_u^T \times f_v \tag{5}$$

3.2. 公平性建模

本文将充分利用用户 - 项目双向图中的数据关系，结合用户自中心图结构层次和节点层次进行公平性建模，将两个层次的目标优化函数融入对抗学习的损失函数中。

3.2.1. 自中心图结构的公平性

构建推荐系统的根本是用户之间的协作性，用户之间通过用户 - 项目双向图结构 \mathcal{G} 中的项目进行关联，所以本节从用户自中心图的角度对公平性进行建模。对于用户 u 而言，经过滤后的用户嵌入表示 f_u 无法全面的表征该用户的局部图结构，因此本文将每个用户 u 建立为以自我为中心的图结构表示：

$$p_u = \mathcal{P}(\mathcal{G}_u, \mathbf{F}) = MLP(h_u^1, h_u^2, \dots, h_u^L) \tag{6}$$

这里使用基于神经网络的方法学习每个用户 u 以自我为中心的第 l 阶邻域表征 h_u^l ，因为多层感知机(MLP)能够逼近任何通用复杂函数，所以使用了 MLP 对子图表示的所有层进行非线性聚合。其中 \mathcal{P} 是用户局部图汇总的结构表示函数， \mathcal{G}_u 是用户自中心图结构。

其中，用户自中心网络的 l 阶邻域在 $l=1$ 和 $l \geq 2$ 分别表示为：

$$h_i^1 = \frac{\sum_{j \in A_i} a_{ij} f_j}{\sum_{j \in A_i} a_{ij}} \quad (7)$$

$$h_i^l = \frac{\sum_{j \in A_i} a_{ij} h_j^{l-1}}{\sum_{j \in A_i} a_{ij}}, \quad \forall l \geq 2 \quad (8)$$

在这里 a_{ij} 表示权重矩阵 \mathbf{A} (公式(1))的边权重， A_i 是节点 i 的相邻节点集。上述公式计算每个节点的第 l 阶自中心图表示，对于 l 的取值会在第五节实验中详细分析。

在得到用户自中心的局部图结构表示 p_u 后，通过对抗训练来确保每个用户的敏感特征不会被局部图结构暴露，本文将自中心图结构层次的公平性目标优化函数定义为：

$$L_S = \mathbb{E}_{(u,v,r,x_u)} \sum_{k=1}^K x_{uk} \ln \mathcal{D}^k(p_u) \quad (9)$$

其中 \mathcal{D}^k 代表鉴别器模块第 k 个鉴别器网络。

3.2.2. 节点层次公平性

平行于用户自中心图结构的表示，过滤敏感特征后向量节点层次的公平性表示也是不容忽视的。节点层次公平性的目标是用户嵌入表示中不会泄露敏感特征，对于给定的敏感特征向量 x_i ，类似于公式(9)，将节点层次的公平性目标优化函数定义为：

$$L_N = \mathbb{E}_{(u,v,r,x_u)} \sum_{k=1}^K x_{uk} \ln \mathcal{D}^k(f_u) \quad (10)$$

其中， f_u 是过滤器模块输出的用户嵌入表示。

公平性要求 L_G 的总目标优化函数可以表示为上述两层公平性表示优化函数的加和，也就是 $L_G = L_S + L_N$ 。

3.3. 对抗训练

对抗训练的主要思想是同时训练过滤器和鉴别器。过滤器的目的是学习推荐任务中的嵌入表示并过滤敏感特征，鉴别器的目的是预测用户敏感特征，弱化过滤器学习敏感特征的能力，二者进行极大极小博弈。具体来说，将过滤后的用户嵌入表示 f_u 作为鉴别器的输入，第 k 个鉴别器 \mathcal{D}^k 负责预测第 k 个敏感特征。因本文重点在建模图结构的公平表示，所以这里的鉴别器网络将使用简单的MLP进行分类预测。对抗训练的过程可以描述为：

$$\arg \max_{\mathcal{F}} \arg \min_{\mathcal{D}} L(\mathcal{F}, \mathcal{D}) = L_R - \lambda L_G = \mathbb{E}_{(u,v,r,x_u)} \left[\ln q_R(r|(f_u, f_v) - \lambda \ln q_{\mathcal{D}}(x|(f_u, p_u))) \right] \quad (11)$$

$$L_R = -\sum_{u=1}^M \sum_{v=1}^V (r_{uv} - \hat{r}_{uv})^2 \quad (12)$$

其中， L_R 是推荐任务的损失函数， L_G 是鉴别器预测属性的损失函数， \hat{r}_{uv} 由公式(5)得出， λ 是对抗系数，其控制着推荐性能和公平性能之间的权重。

4. 实验

在本节中，依次介绍了数据集、评价指标、基线模型和参数设置，然后对模型的性能进行实验研究，继而回答了如下问题：

- RQ1:** 本文提出的 FRC 是否有效提升了推荐公平性?
RQ2: 本文提出的 FRC 应用于基线推荐模型后的推荐性能如何?
RQ3: 用户自中心网络的邻域阶数 l 的取值对模型性能的影响?

4.1. 数据集与采样策略

4.1.1. 数据集

为了研究不同数据规模和稀疏性对推荐性能的影响，本文选择了电影推荐数据集 MovieLens 和电子商务推荐数据集 CIKM。两个数据集的统计结果如表 1 所示。

MovieLens: 该数据集包括 6040 名用户对大约 4000 部电影的 100 万次评级数据，选择性别、年龄和职业作为敏感特征，其中性别为二元特征，职业特征具有 21 个类别，年龄属性则按照用户的年龄范围等长划分为 13 个类别。

CIKM: 该数据集包含 8.8 万个项目和 6 万个用户，210 万个点击实例，选取了性别和年龄作为敏感特征。为了保证数据质量，过滤掉了少于 20 次交互的用户，以形成更密集的数据集。

Table 1. Datasets statistics

表 1. 数据集的统计结果

数据集	Interactions	Users	items	敏感特征
MovieLens	1,000,209	6040	3952	3
CIKM	233,138	1828	57,881	2

4.1.2. 基于曝光的数据负采样

由上节的真实数据集统计信息可知，大量的项目都没有与用户交互过，正样本和负样本项目的数据量悬殊，这也是推荐系统中普遍存在的曝光偏差[11]。所以本文利用基于曝光度的负采样策略来缓解项目曝光偏差，提升推荐质量，保证公平性。首先将项目根据交互次数分为高曝光组 V_{high} 和低曝光组 V_{low} ，这里的交互次数为 1000。由于随机选择的负样本会因为曝光偏差影响推荐质量，所以从低曝光组对负样本进行采样：

$$j \sim \hat{r}_s(u, v), v \in V_{low} \quad (13)$$

4.2. 评估方法

本文使用 NDCG@N 作为评估 Top-N 推荐质量的评价指标，使用 AUC 作为推荐公平性的评估指标。

本文采用了与之前学习公平表征的工作实验设置类似的方法[2] [7] [8]，训练了一组结构与鉴别器完全相同的攻击者。在对抗训练完成后，将过滤后的用户嵌入及其相应的敏感特征输入给攻击者，以训练他们从过滤后的嵌入中对敏感特征进行分类。为每个敏感特征都训练一名攻击者。如果攻击者能够从用户嵌入中准确区分出敏感特征，则认为敏感特征已泄露到用户嵌入中，从而判断推荐模型不公平。

因此检验公平性的本质为分类任务，计算每个攻击者的 AUC 分数，显示攻击者是否能正确对过滤后的用户嵌入进行分类。AUC 得分在范围[0.5, 1]内，越接近 0.5 越好。理想情况下，满足公平性要求的 AUC 得分约为 0.5，表明攻击者无法准确猜测用户嵌入中的敏感特征。公平性的评估指标为 AUC (Area Under the Curve)，其定义如下：

$$AUC = \frac{\sum (t_i, n_j)_{t_i > n_j}}{T * I} \quad (14)$$

其中 T 表示正样本数量, I 表示负样本数量, t_i 表示正样本预测分数, n_j 表示负样本预测分数。

评价推荐任务准确率的常用评价指标 NDCG@N 可以表示为:

$$\text{NDCG}@n = \frac{1}{M} \sum_{i=1}^N \frac{1}{\log_2(p_i + 1)} \quad (15)$$

为了证明提出模型的合理性和有效性, 将 FRC 应用于基础推荐模型 PMF、DMF、BiasedMF 和 DeepModel, 并对比公平推荐模型 FairGAN。各基线模型的具体描述如下:

PMF [12]: 一种浅层概率矩阵分解模型, 在协同过滤场景中应用广泛。

DMF [13]: 一种深度矩阵分解模型, 其使用多层感知机将原始用户 - 项目交互表示为低维向量空间。

BiasedMF [14]: 一种矩阵分解算法, 将用户和物品的偏差加到预测的评分中以适应数据中的偏差和个性化差异。

DeepModel [15]: 该算法通过多层神经网络进行非线性变换以学习用户和项目的复杂表示。

FairGAN [16]: 一种基于 GAN 的推荐模型, 可以公平地为每个项目分配曝光, 并尽可能高的保留用户效用。

4.3. 实验结果与分析(RQ1, RQ2)

表 2 和表 3 展示了在 MovieLens 和 CIKM 数据集上的实验结果, 加粗数据为不同敏感特征中各基线模型中最好的公平性能。前 4 个基线模型为基础推荐模型, FRC_PMF 表示将本文提出的公平推荐模型 FRC 应用于 PMF 模型后的推荐效果。

观察 FRC 与各基线模型的公平性评价指标 AUC 分数得知: 将 FRC 应用于基线模型 PMF、DMF、BiasedMF 和 DeepModel 后的 AUC 分数明显低于表中前 4 个纯基线模型的结果, 这代表 FRC 相较于基础推荐模型在公平推荐性能上有显著的提升。对比基线公平性模型 FairGAN, FRC 的公平性能仍略胜于 FairGAN, 也就是说在这两个数据集上, 敏感特征几乎不可能从过滤后的嵌入表示中被预测出来, 从而保证了公平推荐效果。

观察 FRC 与各基线模型的推荐性能指标 N@5 分数可知: 不论是对比基础推荐模型 PMF、DMF、BiasedMF 和 DMF, 还是公平性基线模型 FairGAN, FRC 在应用于各模型后的仍可以达到较高的推荐性能。一般而言, 保证公平性的同时会牺牲部分推荐性能[2], 因为过滤敏感特征会减少嵌入表示的信息量。因此, 组合敏感特征的推荐性能略低于单个特征是可以理解的。本文提出的 FRC 模型在提升公平性的同时, 仍能使推荐性能接近原始算法, 显示出其在平衡推荐性能和公平性方面的优势, 也代表着利用图对抗学习保证推荐公平的有效性。

Table 2. Compare the performance of the baseline model on the MovieLens dataset. Age, Gen, Occ represent age, gender, and occupation, while Com represents sensitive characteristics. Combination of AUC is the attacker score, and the closer it is to 0.5, the better the fairness performance. N@5 represents NDCG@5

表 2. 在 MovieLens 数据集上和基线模型的性能比较。Age、Gen、Occ 表示年龄、性别和职业, Com.表示敏感特征的组合。AUC 为攻击者分数, 越接近 0.5 表示公平性能越好。N@5 表示 NDCG@5

Model	Age		Gen		Occ		Com.-Age N@5	Com.-Gen N@5	Com.-Occ N@5
	AUC	N@5	AUC	N@5	AUC	N@5			
PMF	0.8428	0.4691	0.7697	0.4819	0.6024	0.4691	0.4846	0.4768	0.4703
DMF	0.7238	0.4281	0.7049	0.3267	0.5710	0.3256	0.3248	0.3298	0.3267
BiasedMF	0.8403	0.4760	0.7870	0.4792	0.6064	0.4757	0.4745	0.4701	0.4672
DeepModel	0.7571	0.3758	0.7165	0.3864	0.5481	0.3816	0.3787	0.3865	0.3833

续表

FairGAN	0.6852	0.4674	0.5982	0.3915	0.6541	0.4562	0.3976	0.4259	0.4523
FRC_PMF	0.5434	0.4672	0.5142	0.4807	0.5223	0.4672	0.4798	0.4718	0.4615
FRC_DMF	0.5241	0.4217	0.5178	0.3212	0.5219	0.3254	0.3217	0.3232	0.3166
FRC_BiasedMF	0.5278	0.4679	0.5027	0.6014	0.5246	0.4714	0.4555	0.4619	0.4615
FRC_DeepModel	0.5356	0.3724	0.5188	0.5384	0.5218	0.3762	0.3569	0.3717	0.3823

Table 3. Performance comparison between baseline model and CIKM dataset**表 3.** 在 CIKM 数据集上和基线模型的性能比较

Model	Age		Gen		Com.
	AUC	N@3	AUC	N@3	N@3
PMF	0.7130	0.6801	0.6654	0.6689	0.6287
DMF	0.7238	0.5141	0.6802	0.4989	0.5173
BiasedMF	0.7123	0.6198	0.7024	0.6029	0.6145
DeepModel	0.6347	0.6354	0.5982	0.6398	0.6289
FairGAN	0.6186	0.6542	0.5615	0.5544	0.6088
FRC_PMF	0.5213	0.6724	0.5122	0.6512	0.6279
FRC_DMF	0.5247	0.4983	0.5167	0.4872	0.5123
FRC_BiasedMF	0.5391	0.6072	0.5243	0.5709	0.6089
FRC_DeepModel	0.5166	0.6289	0.5199	0.6394	0.6173

4.4. 参数分析(RQ3)

本节研究了以用户为中心的子图网络邻域阶数 l 的取值对 FRC 模型性能的影响。表 4 展示了不同 l 的取值对模型性能的数据影响，图 2 为表 4 数据的直观折线图。观察图表数据可知，高阶图结构可以得到更好的公平性结果和更高的推荐准确性，但对高阶图进行建模也会使模型训练更加耗时。因 l 的取值大于 3 时用户的子图与其他用户存在较大的重叠，故仅研究 $l \leq 3$ 的情况。其他的数据集、评价指标和基线模型的组合均观察到了类似的数据趋势，故这里仅展示了 FRC_BiasedMF 模型在 MovieLens 数据集上单个敏感特征的实验数据。

Table 4. The influence of the value of the neighborhood order l of the user's self-centered network on FRC**表 4.** 用户自中心网络的邻域阶数 l 的取值对 FRC 的影响

敏感特征	$l = 1$		$l = 2$		$l = 3$	
	AUC	N@5	AUC	N@5	AUC	N@5
Age	0.5431	0.4793	0.5344	0.4741	0.5311	0.4723
Gen	0.5292	0.4962	0.5267	0.4952	0.5221	0.4921
Occ	0.5341	0.4897	0.5263	0.4867	0.5258	0.4862

5. 结论

目前，推荐系统因数据偏差的影响而普遍存在公平性问题。本文提出 FRC 模型通过构建用户 - 项目

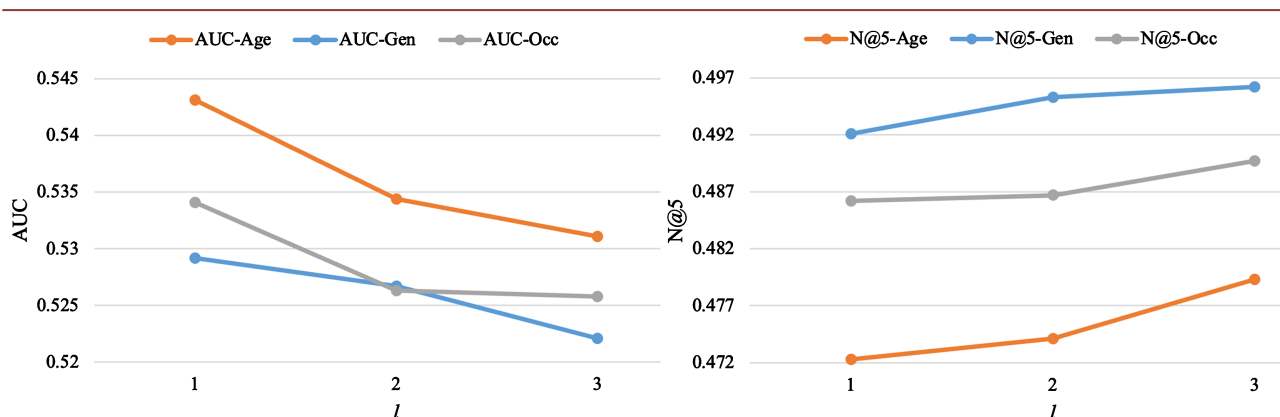


Figure 2. The influence of subgraph neighborhood order l on model performance

图 2. 子图邻域阶数 l 对于模型性能的影响

双向图并利用图对抗技术学习公平嵌入，确保推荐结果的公平性，经真实数据集验证了该模型的有效性，未来我们将进一步探索 FRC 模型在教育和工作推荐等领域的应用潜力。

基金项目

国家自然科学基金项目(61772342)资助。

参考文献

- [1] Geyik, S.C., Ambler, S., and Kenthapadi, K. (2019) Fairness-Aware Ranking in Search & Recommendation Systems with Application to LinkedIn Talent Search. *KDD'19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage, 4-8 August 2019, 2221-2231. <https://doi.org/10.1145/3292500.3330691>
- [2] Bose, A. and Hamilton, W. (2019) Compositional Fairness Constraints for Graph Embeddings. <https://doi.org/10.48550/arXiv.1905.10674>
- [3] Zhu, Z., Hu, X. and Caverlee, J. (2018) Fairness-Aware Tensor-Based Recommendation. *Proceedings of CIKM '18*, Torino, 1153-1162.
- [4] Datta, A., Tschantz, M.C. and Datta, A. (2015) Automated Experiments on AD Privacy Settings: A Tale of Opacity, Choice, and Discrimination. *Proceedings on Privacy Enhancing Technologies*, **1**, 92-112. <https://doi.org/10.1515/popets-2015-0007>
- [5] Karimi, F., Génois, M., Wagner, C., Singer, P. and Strohmaier, M. (2018) Homophily Influences Ranking of Minorities in Social Networks. *Scientific Reports*, **8**, 1-12. <https://doi.org/10.1038/s41598-018-29405-7>
- [6] Wu, L., Sun, P., Fu, Y., Hong, R., et al. (2019) A Neural Influence Diffusion Model for Social Recommendation. *SIGIR'19: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Paris, 21-25 July 2019, 235-244. <https://doi.org/10.1145/3331184.3331214>
- [7] Wu, L., Yang, Y., Zhang, K., Hong, R., et al. (2020) Joint Item Recommendation and Attribute Inference: An Adaptive Graph Convolutional Network Approach. *SIGIR'20: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Xi'an, 25-30 July 2020, 679-688. <https://doi.org/10.1145/3397271.3401144>
- [8] Chen, L., Wu, L., Hong, R., et al. (2020) Revisiting Graph based Collaborative Filtering: A Linear Residual Graph Convolutional Network Approach. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, AAAI-20 Technical Tracks 1. <https://doi.org/10.1609/aaai.v34i01.5330>
- [9] Zemel, R., Wu, Y., Swersky, K., Pitassi, T. and Dwork, C. (2013) Learning Fair Representations. *ICML*, **2010**, 325-333.
- [10] Li, Y., Chen, H., Xu, S., Ge, Y. and Zhang, Y. (2021) Towards Personalized Fairness Based on Causal Notion. *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21)*, New York, 11-15 July 2021, 1-10.
- [11] Chen, J., Dong, H., Wang, X., et al. (2020) Bias and Debias in Recommender System: A Survey and Future Directions.

- <https://doi.org/10.48550/arXiv.2010.03240>
- [12] Salakhutdinov, R. and Mnih, A. (2007) Probabilistic Matrix Factorization. *Proceedings of the 20th International Conference on Neural Information Processing Systems (NIPS'07)*, Red Hook, 1257-1264.
 - [13] Xue, H.-J., Dai, X., Zhang, J., *et al.* (2017) Deep Matrix Factorization Models for Recommender Systems. <https://doi.org/10.1016/j.ecoinf.2022.101775>
 - [14] Koren, Y., Bell, R. and Volinsky, C. (2009) Matrix Factorization Techniques for Recommender Systems. *Computer*, **8**, 30-37. <https://doi.org/10.1109/MC.2009.263>
 - [15] Cheng, H.-T., Koc, L., Harmsen, J., Shaked, T., *et al.* (2016) Wide & Deep Learning for Recommender Systems. *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, Boston, 15 September 2016, 7-10. <https://doi.org/10.1145/2988450.2988454>
 - [16] Li, J., Ren, Y., Deng, K., *et al.* (2022) FairGAN: GANs-Based Fairness-Aware Learning for Recommendations with Implicit Feedback. *WWW'22: Proceedings of the ACM Web Conference 2022*, Lyon, 25-29 April 2022, 297-307. <https://doi.org/10.1145/3485447.3511958>