Selective Fairness in Recommendation via Prompts

Yiqing Wu[*]
Institute of Computing Technology,
Chinese Academic of Science;
University of Chinese Academy of
Sciences;
WeChat Search Application
Department, Tencent, China
wuyiqing20s@ict.ac.cn

Ruobing Xie[†]
WeChat Search Application
Department, Tencent, China
ruobingxie@tencent.com

Yongchun Zhu[*]
Institute of Computing Technology,
Chinese Academic of Science;
University of Chinese Academy of
Sciences, China
zhuyongchun18s@ict.ac.cn

Fuzhen Zhuang[‡]
Institute of Artificial Intelligence,
Beihang University, Beijing;
SKLSDE, School of Computer Science,
Beihang University, China
zhuangfuzhen@buaa.edu.cn

Ao Xiang[*]
Institute of Computing Technology,
Chinese Academic of Science;
University of Chinese Academy of
Sciences, China
aoxiang@ict.ac.cn

Xu Zhang
WeChat Search Application
Department, Tencent, China
xuonezhang@tencent.com

Leyu Lin
WeChat Search Application
Department, Tencent, China
goshawklin@tencent.com

Qing He[*][‡]
Institute of Computing Technology,
Chinese Academic of Science;
University of Chinese Academy of
Sciences, China
heqing@ict.ac.cn

ABSTRACT

Recommendation fairness has attracted great attention recently. In real-world systems, users usually have multiple sensitive attributes (e.g. age, gender, and occupation), and users may not want their recommendation results influenced by those attributes. Moreover, which of and when these user attributes should be considered in fairness-aware modeling should depend on users' specific demands. In this work, we define the selective fairness task, where users can flexibly choose which sensitive attributes should the recommendation model be bias-free. We propose a novel parameter-efficient prompt-based fairness-aware recommendation (PFRec) framework, which relies on attribute-specific prompt-based bias eliminators with adversarial training, enabling selective fairness with different attribute combinations on sequential recommendation. Both task-specific and user-specific prompts are considered. We conduct extensive evaluations to verify PFRec's superiority in selective fairness. The source codes are released in https://github.com/wyqing20/PFRec.

[*]The authors are at the Key Lab of Intelliegent Information Processing of Chinese Academy of Sciences (CAS). Xiang Ao is alao at the Institute of Intelligent Computing Technology, Suzhou, CAS.
[†]indicates equal contributions.
[‡]indicates corresponding author.

CCS CONCEPTS

• **Information systems** → **Recommender systems**.

KEYWORDS

recommendation, prompt, fairness

1 INTRODUCTION

With the exponential growth of information, recommender systems become more and more essential for us to efficiently get useful information. The blooming personalized recommendation algorithms facilitate providing more attractive items, while these data-driven learning and objectives inevitably bring in unfairness [5, 25].

The *recommendation fairness* has attracted broad attention in recent years, which contains multisided factors from consumers, providers, and both [4]. To achieve the consumer-side fairness, a general idea is to make the learned user preferences or the recommended results bias-free on users' sensitive attributes (e.g., gender, age, country, income level) [14]. Various of techniques have been adopted for fairness modeling, such as Pareto optimization [15], adversarial training [21], and graph-based models [22].

Building a high-quality fairness-aware recommender system often involves with compromises, since user sensitive attributes

are also essential sources of personalization. Eliminating too many user attributes is likely to harm the recommendation accuracy [21]. Moreover, in complicated real-world recommender systems, there are multiple types of user sensitive attributes. Users' preferences on which of and when these sensitive attributes should be eliminated are also personalized according to their practical demands. For example, the attribute *age* is often a good feature for the majority of users that helps algorithms to recommend popular songs among their peers, while sometimes other users do want to escape from the information cocoon and explore more possibilities in music. It will be fantastic if users can have flexible **attribute-specific fairness switches** that could customize which user sensitive attributes should be eliminated in recommended results at this time.

In this work, we define the **selective fairness** in practical recommendation, which aims to give users the flexibility to choose and change which sensitive attributes should be (temporally) ignored in sequential recommendation. Enabling such selective fairness mainly has two challenges: (1) the number of attribute combinations (2^m) grows exponentially as the number of attributes (m) increases, while most conventional fairness-aware models design specific models/parameters for each setting [15, 21]. It is extremely inefficient (and even impossible) to fully train and store different fairness-aware models for all 2^m attribute combinations. (2) Data sparsity is a widespread fundamental issue in recommendation. The proposed selective fairness models for all fairness settings should make full use of the shared behavioral information that has not been debiased to balance recommendation accuracy and fairness.

To address these problems, we propose a novel **Prompt-based fairness-aware recommendation (PFRec)**, which introduces the popular and widely-verified parameter-efficient techniques from NLP to sequential recommendation. A prompt is often a small piece of hard texts or soft embeddings, which helps to effectively extract useful information from the huge pre-training models to serve for the downstream tasks at a small tuning cost [3, 12]. Specifically in PFRec, we first train the original sequential recommendation model on all user historical behaviors as the pre-training model, which does not consider any type of fairness. Second, inspired by prompts, we build a set of **prompt-based bias eliminators** with *personalized attribute-specific prompts* and *adapters* specially designed to ensure fairness on different attribute combinations in the learned user representations. During the prompt-tuning, only the bias eliminators are updated with the pre-trained sequential model fixed. Finally, a generative adversarial network (GAN) is adopted, where the discriminator aims to distinguish whether the prompt-enhanced user representations have biases in the selected attribute combination. The advantages of our PFRec are as follows: (1) the parameter-efficient prompt-tuning and adapters balance effectiveness and efficiency, which makes the selective fairness models more feasible in practical systems. (2) The pre-training + prompt-tuning framework takes full advantage of the pre-training model's modeling ability learned from all user historical behaviors, maintaining good recommendation ability while considering selective fairness.

In experiments, we conduct extensive evaluations on two public datasets with both single-attribute and multi-attribute fairness settings. The results verify the effectiveness of PFRec in jointly considering accuracy and attribute-specific fairness. The contributions of this work are concluded as follows:

- We highlight the challenges of selective fairness in sequential recommendation, and design a novel Prompt-based fairness-aware recommendation to address this task. To the best of our knowledge, we are the first to adopt prompts in recommendation for fair sequence modeling.
- We propose a personalized attribute-specific prompt to efficiently and flexibly fit for different attribute combinations in fairness-aware recommendation.
- We have verified the effectiveness of PFRec on multiple datasets with different attribute combinations.

2 METHODOLOGY

2.1 Preliminaries

Background of Prompt. A prompt is often a small piece of hard texts or soft embeddings inserted into the original sequence, which helps to efficiently extract knowledge from the pre-training models for the downstream tasks [3, 13]. During tuning, only the prompts (having much fewer parameters) will be updated, with the whole pre-trained model unchanged. In this case, we do not need to store different huge language models tuned for every downstream task. Hence, prompt is **parameter-efficient** especially when there are lots of downstream tasks (e.g., our selective fairness task).

Our Task of Selective Fairness. In this work, we focus on the selective fairness task on sequential recommendation. In real-world systems, there are multiple user sensitive attributes that could be considered in fairness-aware modeling, such as age, gender, country, and other user profiles. However, different users (or even a user at different time) may have different demands on which attributes' fairness should be considered. It is unaffordable to train different separate models for all 2^m attribute combinations. The selective fairness task aims to build a general parameter-efficient framework, which can provide attribute-specific fairness for different combinations of user attributes flexibly selected by users.

Notions of PFRec. Formally, we denote user and item as $u \in U$ and $v \in V$, where U and V are the overall user and item set. Each user u has a historical behavior sequence $s_u = \{v_1^u, v_2^u, ..., v_{|s_u|}^u\}$ (of length $|s_u|$) and m user sensitive attributes $A_u = \{a_1^u, a_2^u, ..., a_m^u\}$ (making 2^m attribute combinations). Each user can select a set of attributes to be considered for fairness. In pre-training, we first train a general sequential recommendation model $f_{seq}(s_u|\Theta)$ on all user behaviors, where $f_{seq}(.)$ is the sequential model and Θ is its parameters. In tuning, we build the prompt-based bias eliminator for the k-th attribute combination, and the attribute-specific fairness-aware sequence modeling is noted as $f_{seq}(s_u|\Theta, \vartheta^k)$, where ϑ^k is the parameters of the k-th parameter-efficient eliminator to be updated in tuning. Note that PFRec has 2^m "attribute-specific" fairness settings related to the number of total *attribute combinations*.

2.2 Overall Framework

The overall framework of PFRec is shown in Fig. 1. We first train a general sequential model based on all user historical behaviors as the pre-training model, which does not intentionally model any type of fairness. In tuning, we adopt a GAN framework to ensure the fairness on different attribute combinations. The discriminator aims to distinguish whether the learned user representations are biased on selected attributes, while the generator (i.e., the sequential

model) aims to (a) recommend appropriate items and (b) generate unbiased user representations on selected attributes. To enhance the fairness of the generator, for a certain attribute combination, we build an attribute-specific prompt-based bias eliminator to purify the attribute biases in learned user representations. This eliminator contains a prefix soft personalized prompt inserted into the behavior sequence, which is built from both the attribute combination to be debiased and the specific user attributes of u. Moreover, we also adopt the adapter technique [8] as a supplement for bias elimination. The prompt-enhanced sequence is then fed into the pre-trained sequential model and output a bias-free user representation. During tuning, only the prompt-based bias eliminator will be updated to enable a parameter-efficient fairness-aware tuning.
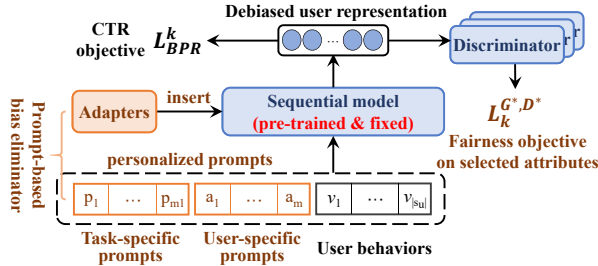

Figure 1: Overall architecture of PFRec.

2.3 Prompt-based Bias Eliminator

2.3.1 Pre-training. The pre-training model is the original base model with no fairness consideration. Without loss of generality, we adopt a classical SASRec model [11] as the sequential model in PFRec, which adopts the powerful self-attention layers for behavior modeling. For a input sequence s_u, we define its l-layer's behavior representation matrix as $H_u^l = \{h_{u,1}^l, h_{u,2}^l, \cdots, h_{u,|s_u|}^l\}$, where $h_{u,i}^l$ is the i-th behavior's representation of u at the l-th layer. The final user representation u is pre-training is learned as:

$$u = f_{seq}(s_u|\Theta) = h_{u,|s_u|}^L, \quad H_u^{l+1} = \text{Transformer}^l(H_u^l). \quad (1)$$

Θ is the sequential model's parameters updated in pre-training.

2.3.2 Attribute-specific Bias Eliminator in Tuning. After learning the general sequential model in pre-training, PFRec aims to build attribute-specific fairness-aware models in tuning via prompt-based bias eliminator. Specifically, it constructs a prefix soft personalized prompt, which contains a task-specific part and a user-specific part. The task-specific prompt is related to the specific attribute combination to be considered in fairness debasing, which aims to tell the eliminator which attributes it should focus on. We use $m_1 = 10$ randomly initialized tokens as the task-specific prompt p_t^k for the k-th attribute combination. As for the personalization information, we directly use the m attribute embeddings of u to form the user-specific prompt $p_u = \{a_1^u, a_2^u, \cdots, a_m^u\}$. These two prompts are simply added in front of the user behaviors to get a new prompt-enhanced sequence as $\hat{s}_u^k = \{p_t^k, p_u, v_1^u, v_2^u, ..., v_{|s_u|}^u\}$. The corresponding l-layer behavior matrix is noted as $\hat{H}_u^{l,k}$.

To further improve the fairness-oriented debiasing, we insert adaptors into the Transformer block, which is also verified as an effective parameter-efficient tuning technique [8]. Specifically, we insert adaptors after the multi-head and FFN modules, noted as:

$$\text{Transformer}_a^l(\hat{H}_u^{l,k}) = \text{Adapter}_1^{l,k}(\text{FFN}(\text{Adapter}_2^{l,k}(\text{ATT}(\hat{H}_u^{l,k})))),$$

$$\text{Adapter}^k(X) = \text{LayerNorm}(XW_d^k)W_u^k + X, \quad (2)$$

in which W_d^k and W_u^k are trainable matrices for the k-th attribute combination. Finally, the fairness-enhanced user representation $\hat{u}^k$ for the k-th combination is learned on $\hat{s}_u^k$ via $\text{Transformer}_a(\cdot)$ as:

$$\hat{u}^k = f_{seq}(\hat{s}_u^k|\Theta, \vartheta^k) = \hat{h}_{u,m_1+m+|s_u|}^{L,k}. \quad (3)$$

The user embedding $\hat{u}^k$ is supposed to be bias-free on the selected attributes of the k-th combination, while maintaining the ability of personalized recommendation learned in pre-training.

2.4 Adversarial Training Objectives

To filter attribute biases in $\hat{u}^k$ while remaining personalization for recommendation, we rely on the adversarial training. Specifically, the eliminator-enhanced sequential model is regarded as the generator to build $\hat{u}^k$, while the discriminator is a set of classifiers to predict u's actual labels on the selected attributes.

Following the classical GAN's minimax game [6], the adversarial objective $L_k^{G^*, D^*}$ for the k-th attribute combination is defined as:

$$L_k^{G^*, D^*} = \min_{\phi^k} \max_{\vartheta^k} \sum_B \sum_{u \in B} \sum_{a_i \in Q_k} -\mathbb{E}_{\hat{u}^k}[\log(P(a_i|\hat{u}^k, \phi^k))]. \quad (4)$$

Here, $P(a_i|\hat{u}^k, \phi^k)$ indicates the probability of the discriminator having predicted the correct attribute of a_i for u based on $\hat{u}^k$ learned from the generator. ϕ^k is the parameter set for the discriminator corresponding to the k-th attribute combination. Q_k is the selected attribute set of the k-th setting, B is the batch. Note that the generator only updates the prompt-based eliminator's parameters ϑ^k.

To maintain the ability of personalized recommendation in the generator, we further adopt a classical BPR loss [11, 18] for the generator as L_k^{BPR} to update ϑ^k. The overall loss of PFRec L_k on the k-th attribute combination is formalized as follows:

$$L_k = L_{BPR}^k + \lambda L_k^{G^*, D^*}. \quad (5)$$

λ is a loss weight empirically set to be 1. Through the adversarial training in Eq. (5) via attribute-specific prompt-based bias eliminators, all 2^m tuned fairness-aware models for different attribute combinations could be learned properly.

3 EXPERIMENTS

In this section, we aim at answering the following research questions: (RQ1) How does our PFRec perform compared with conventional models and fairness models in considering single-attribute fairness on various evaluation metrics? (RQ2) How does our PFRec perform compared with conventional models and fairness models in considering single-attribute fairness? (RQ3) How does PFRec perform compared with its various ablation versions?

3.1 Experimental Settings

Dataset. We evaluate PFRec on two real-world open datasets, including CIKM [34] and AliEC. The CIKM dataset is an E-commerce recommendation dataset[1]. It has 88 thousand items and 60 thousand users with 2.1 million click instances. Each user has 3 attributes: gender, age, consumption level. The AliEC dataset[2] contains nearly 100 thousand users and 8.8 million click instances. Each user has two attributes gender and age. For each dataset, we filter users having less than 10 instances. We adopt the leave-one-out strategy to split datasets, using the most recent item of each user for testing and the second most recent item for validation.

Competitors. In this work, we use the classical SASRec [11] as the base sequential model used in pre-training, while it is also convenient to deploy PFRec on other models (e.g., BERT4Rec [20]). We compare PFRec with several baselines: (1) **SASRec**, which is the pre-training model without considering any fairness. (2) **BERT4Rec [20]**, which is a strong BERT-based sequential model. (3) **Li et al. [14]***. This work is a SOTA fairness model for general recommendation. We follow its separate method, build an attribute-specific filter after u learned by SASRec, and train it with the same loss in Eq. (5). All models share the same behaviors and features if needed.

Evaluation Protocols and Parameter Settings. For accuracy, we use the classical AUC, HIT@N, and NDCG@N as our metrics with $N = 10$. For each test instance, we randomly sample 99 items that the user did not click as negative samples as [20, 32]. For fairness, following the settings in [14, 21], we train a set of attackers which have the same structure and capacity as the discriminators to evaluate selective fairness. Following [21], we adopt micro-averaged F1 to measure the fairness performance. The smaller values of F1 denote better fairness performances of models. For parameters, the embedding size is 64 and batch size is 256 equally for all methods. The loss weight $\lambda = 1.0$. We train the base sequential model and attackers by Adam optimizer. Following [1], we optimize the generator and discriminator by RMSprop optimizer. The learning rate of our model is set to $1e - 4$, and the L2 normalization coefficient is set as $1e - 6$. We conduct a grid search for hyper-parameters.

3.2 Single-attribute Evaluation (RQ1)

In single-attribute evaluation, we require fairness on only one attribute. Table 1 and Table 2 show the results of all single attributes. We can find that: (1) PFRec achieves the best fairness performances on all attributes in two datasets. The improvements are significant ($p < 0.05$). Compared with the base model SASRec, PFRec has over $12\% - 41\%$ relative improvements on the fairness metrics of all sensitive attributes. It indicates that our prompt-based bias eliminator is effective especially for sequential behaviors. (2) PFRec outperforms Li et al. [14]* in almost all accuracy metrics. It is because that our prompt-based bias eliminators are perfectly suitable for extracting useful personalized but not attribute-biased user preferences in sequential recommendation. Although it is natural that PFRec performs slightly worse than the base model, while the overall performance is still acceptable considering the fairness gains.

3.3 Multi-attribute Evaluation (RQ2)

We also report the results of multi-attribute fairness in Table 3. Due to the limited space, we directly give the fairness setting where all attributes are considered. We observe that: PFRec still significantly outperforms baselines on most fairness metrics, especially on AliEC. It verifies that our attribute-specific bias eliminators can cooperate well on joint fairness demands. It also implies the power of our parameter-efficient techniques in fairness-aware tuning.

3.4 Ablation Study (RQ3)

We further conduct an ablation study. Fig. 2 shows the relative fairness (F1) and accuracy (AUC) improvements over the pre-training model on AliEC. We find that: (1) PFRec outperforms PFRec(w/o Prompt), while all PFRec models perform significantly better than pre-train in fairness. It verifies the effectiveness of both prompts and adapters in bias eliminators. (2) The fine-tuning model is a parameter-inefficient method which updates all parameters including Θ via Eq. (5). Although fine-tuning has comparable accuracy and fairness performances with PFRec, it is impossible to deploy it for all attribute combinations and store 2^m fully tuned models. Our PFRec is still the best choice to address the selective fairness issue in practical sequential recommender systems.
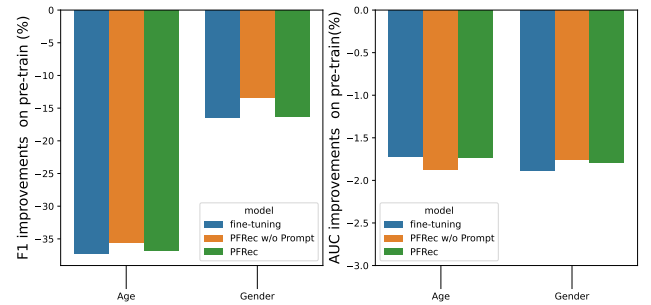

Figure 2: Ablation study on (a) fairness↓ and (b) accuracy↑.

4 RELATED WORKS

Fairness in Recommendation. Recommendation fairness is measured from different views [2, 5, 17, 21, 22]. Some of them solve fairness from the provider side [10, 16], making items from different providers be treated equally. Other works focus on the customer view via group-based [21, 25] and individual-based [5] methods. Various techniques such as Pareto efficiency [15, 25] and adversarial training [21] are used for unbiased preference learning. Li et al. [14] focuses on personalized fairness in general recommendation based on causal notion. Different from these models, we focus on the selective fairness in sequential recommendation.

Sequential Recommendation. User historical behaviors are essential information to reflect user preferences. There are lots of sequential models verified in recommendation [7, 9, 31, 33]. Recently, the self-attention based encoder has been widely adopted for sequence modeling [11, 20, 26] and feature interaction [19, 28, 29]. Self-supervised learning and pre-training are also used for further improvements in recommendation [23, 24, 27, 30, 32].

Table 1: Fairness (F1) and Accuracy (AUC, H@10, N@10) results of single-attribute evaluation on CIKM.

Model	Gender				Age				Consumption level			
	F1↓	AUC↑	H@10↑	N@10↑	F1↓	AUC↑	H@10↑	N@10↑	F1↓	AUC↑	H@10↑	N@10↑
SASRec	0.8998	0.8857	0.7092	0.5444	0.4491	0.8857	0.7092	0.5444	0.1900	0.8857	0.7092	0.5444
BERT4Rec	0.9104	0.8713	0.6625	0.4913	0.4931	0.8713	0.6625	0.4913	0.2114	0.8713	0.6625	0.4913
Li et al. [14]*	0.7975	0.8386	0.5830	0.3830	0.2818	0.8396	0.5889	0.3880	0.1414	0.8666	0.6478	0.4380
PFRec	**0.7843**	0.8504	0.6319	0.4748	**0.2661**	0.8525	0.6384	0.4827	**0.1289**	0.8709	0.6799	0.5176

Table 2: Fairness (F1) and Accuracy (AUC, H@10, N@10) results of single-attribute evaluation on AliEC.

Model	Age				Gender			
	F1↓	AUC↑	H@10↑	N@10↑	F1↓	AUC↑	H@10↑	N@10↑
SASRec	0.4310	0.9060	0.7499	0.5412	0.8444	0.9060	0.7499	0.5412
BERT4Rec	0.4420	0.8923	0.6993	0.4754	0.8623	0.8923	0.6993	0.4754
Li et al. [14]*	0.3978	0.8898	0.6995	0.4756	0.8353	0.8893	0.6992	0.4785
PFRec	**0.2774**	0.8890	0.7042	0.4938	**0.7315**	0.8901	0.7064	0.4948

Table 3: Fairness (F1) and Accuracy (AUC, H@10, N@10) results of multi-attribute evaluation on CIKM and AliEC.

Model	CIKM						AliEC				
	F1-G↓	F1-A↓	F1-C↓	AUC↑	H@10↑	N@10↑	F1-G↓	F1-A↓	AUC↑	H@10↑	N@10↑
SASRec	0.8998	0.4491	0.1900	0.8857	0.7092	0.5444	0.8444	0.4310	0.9060	0.7499	0.5412
BERT4Rec	0.9104	0.4931	0.2114	0.8713	0.6625	0.4913	0.8623	0.4420	0.8923	0.6993	0.4754
Li et al. [14]*	**0.7879**	0.2532	0.1114	0.8386	0.5295	0.3439	0.8122	0.3803	0.8765	0.6621	0.4377
PFRec	0.7903	**0.2418**	**0.1057**	0.8504	0.5638	0.4194	**0.6872**	**0.2661**	0.8762	0.6688	0.4588

5 CONCLUSION AND FUTURE WORK

In this work, considering different recommendation fairness demands of different users, we propose PFRec for enabling selective fairness in practical systems. Specifically, we apply prompt-based attribute-specific bias eliminators to the sequential inputs and the inner structures of sequential model to debias the user representation according to users' demands. The extensive evaluations verify the effectiveness of PFRec in modeling selective fairness while maintaining recommendation accuracy. In the future, we will explore more pre-training/prompt-tuning tasks to further improve the performances.

ACKNOWLEDGMENTS

REFERENCES

[1] Martin Arjovsky and Léon Bottou. 2017. Towards principled methods for training generative adversarial networks. *arXiv preprint arXiv:1701.04862* (2017).
[2] Alex Beutel, Jilin Chen, Tulsee Doshi, Hai Qian, Li Wei, Yi Wu, Lukasz Heldt, Zhe Zhao, Lichan Hong, Ed H Chi, et al. 2019. Fairness in recommendation ranking through pairwise comparisons. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2212–2220.
[3] Tom B Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *arXiv preprin* (2020).
[4] Robin Burke. 2017. Multisided fairness for recommendation. *arXiv preprint arXiv:1707.00093* (2017).

[5] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*. 214–226.
[6] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Proceedings of NIPS*.
[7] Xiaobo Hao, Yudan Liu, Ruobing Xie, Kaikai Ge, Linyao Tang, Xu Zhang, and Leyu Lin. 2021. Adversarial Feature Translation for Multi-domain Recommendation. In *Proceedings of KDD*.
[8] Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin De Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. 2019. Parameter-efficient transfer learning for NLP. In *International Conference on Machine Learning*. PMLR, 2790–2799.
[9] Dietmar Jannach and Malte Ludewig. 2017. When recurrent neural networks meet the neighborhood for session-based recommendation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*. 306–310.
[10] Toshihiro Kamishima, Shotaro Akaho, Hideki Asoh, and Jun Sakuma. 2014. Correcting Popularity Bias by Enhancing Recommendation Neutrality.. In *RecSys Posters*.
[11] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *Proceedings of ICDM*.
[12] Brian Lester, Rami Al-Rfou, and Noah Constant. 2021. The power of scale for parameter-efficient prompt tuning. In *Proceedings of EMNLP*.
[13] Xiang Lisa Li and Percy Liang. 2021. Prefix-tuning: Optimizing continuous prompts for generation. *ACL-IJNLP* (2021).
[14] Yunqi Li, Hanxiong Chen, Shuyuan Xu, Yingqiang Ge, and Yongfeng Zhang. 2021. Towards personalized fairness based on causal notion. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1054–1063.
[15] Xiao Lin, Hongjie Chen, Changhua Pei, Fei Sun, Xuanji Xiao, Hanxiao Sun, Yongfeng Zhang, Wenwu Ou, and Peng Jiang. 2019. A pareto-efficient algorithm for multiple objective optimization in e-commerce recommendation. In *Proceedings of the 13th ACM Conference on recommender systems*. 20–28.
[16] Weiwen Liu, Jun Guo, Nasim Sonboli, Robin Burke, and Shengyu Zhang. 2019. Personalized fairness-aware re-ranking for microlending. In *Proceedings of the 13th ACM Conference on Recommender Systems*. 467–471.
[17] Evaggelia Pitoura, Kostas Stefanidis, and Georgia Koutrika. 2021. Fairness in rankings and recommendations: An overview. *The VLDB Journal* (2021), 1–28.

[18] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. *UAI* (2009).

[19] Weiping Song, Chence Shi, Zhiping Xiao, Zhijian Duan, Yewen Xu, Ming Zhang, and Jian Tang. 2019. Autoint: Automatic feature interaction learning via self-attentive neural networks. In *Proceedings of CIKM*.

[20] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. In *Proceedings of CIKM*.

[21] Chuhan Wu, Fangzhao Wu, Xiting Wang, Yongfeng Huang, and Xing Xie. 2021. Fairrec: fairness-aware news recommendation with decomposed adversarial learning. AAAI.

[22] Le Wu, Lei Chen, Pengyang Shao, Richang Hong, Xiting Wang, and Meng Wang. 2021. Learning fair representations for recommendation: A graph-based perspective. In *Proceedings of the Web Conference 2021*. 2198–2208.

[23] Yiqing Wu, Ruobing Xie, Yongchun Zhu, Xiang Ao, Xin Chen, Xu Zhang, Fuzhen Zhuang, Leyu Lin, and Qing He. 2022. Multi-view Multi-behavior Contrastive Learning in Recommendation. In *Proceedings of DASFAA*.

[24] Chaojun Xiao, Ruobing Xie, Yuan Yao, Zhiyuan Liu, Maosong Sun, Xu Zhang, and Leyu Lin. 2021. UPRec: User-Aware Pre-training for Recommender Systems. *arXiv preprint* (2021).

[25] Lin Xiao, Zhang Min, Zhang Yongfeng, Gu Zhaoquan, Liu Yiqun, and Ma Shaoping. 2017. Fairness-aware group recommendation with pareto-efficiency. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*. 107–115.

[26] Ruobing Xie, Cheng Ling, Yalong Wang, Rui Wang, Feng Xia, and Leyu Lin. 2020. Deep Feedback Network for Recommendation. In *Proceedings of IJCAI*.

[27] Ruobing Xie, Qi Liu, Liangdong Wang, Shukai Liu, Bo Zhang, and Leyu Lin. 2021. Contrastive Cross-domain Recommendation in Matching. *arXiv preprint arXiv:2112.00999* (2021).

[28] Ruobing Xie, Zhijie Qiu, Jun Rao, Yi Liu, Bo Zhang, and Leyu Lin. 2020. Internal and Contextual Attention Network for Cold-start Multi-channel Matching in Recommendation. In *Proceedings of IJCAI*.

[29] Ruobing Xie, Shaoliang Zhang, Rui Wang, Feng Xia, and Leyu Lin. 2021. Hierarchical Reinforcement Learning for Integrated Recommendation. In *Proceedings of AAAI*.

[30] Zheni Zeng, Chaojun Xiao, Yuan Yao, Ruobing Xie, Zhiyuan Liu, Fen Lin, Leyu Lin, and Maosong Sun. 2021. Knowledge transfer via pre-training for recommendation: A review and prospect. *Frontiers in big Data* (2021).

[31] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. In *Proceedings of KDD*.

[32] Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. S3-rec: Self-supervised learning for sequential recommendation with mutual information maximization. In *CIKM*.

[33] Yongchun Zhu, Dongbo Xi, Bowen Song, Fuzhen Zhuang, Shuai Chen, Xi Gu, and Qing He. 2020. Modeling users' behavior sequences with hierarchical explainable network for cross-domain fraud detection. In *Proceedings of The Web Conference 2020*.

[34] Yongchun Zhu, Ruobing Xie, Fuzhen Zhuang, Kaikai Ge, Ying Sun, Xu Zhang, Leyu Lin, and Juan Cao. 2021. Learning to warm up cold item embeddings for cold-start recommendation with meta scaling and shifting networks. In *Proceedings of SIGIR*. 1167–1176.