Personalized Transfer of User Preferences for Cross-domain Recommendation

Yongchun Zhu[1,3,4], Zhenwei Tang[1], Yudan Liu[4], Fuzhen Zhuang[2,5,*], Ruobing Xie[4], Xu Zhang[4], Leyu Lin[4] and Qing He[1,3]
[1]Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China
[2]Institute of Artificial Intelligence, Beihang University, Beijing 100191, China
[3]University of Chinese Academy of Sciences, Beijing 100049, China
[4]WeChat Search Application Department, Tencent, China
[5]SKLSDE, School of Computer Science, Beihang University, Beijing 100191, China
{zhuyongchun18s, heqing}@ict.ac.cn, lilvjosephtang@gmail.com, {danydliu, ruobingxie, xuonezhang, goshawklin}@tencent.com,zhuangfuzhen@buaa.edu.cn

（WSDM-2022) https://github.com/easezyc/WSDM2022-PTUPCDR

gesis Leibniz-Institut für Sozialwissenschaften

Forschungszentrum · Research Center L3S

Reported by Ke Gan

1. **Introduction**
2. **Approach**
3. **Experiments**

Introduction

Figure 1: (a) In existing CDR methods: all users share the common bridge function. (b) The proposed PTUPCDR utilizes a meta network to generate personalized bridge functions for each user.

Approach

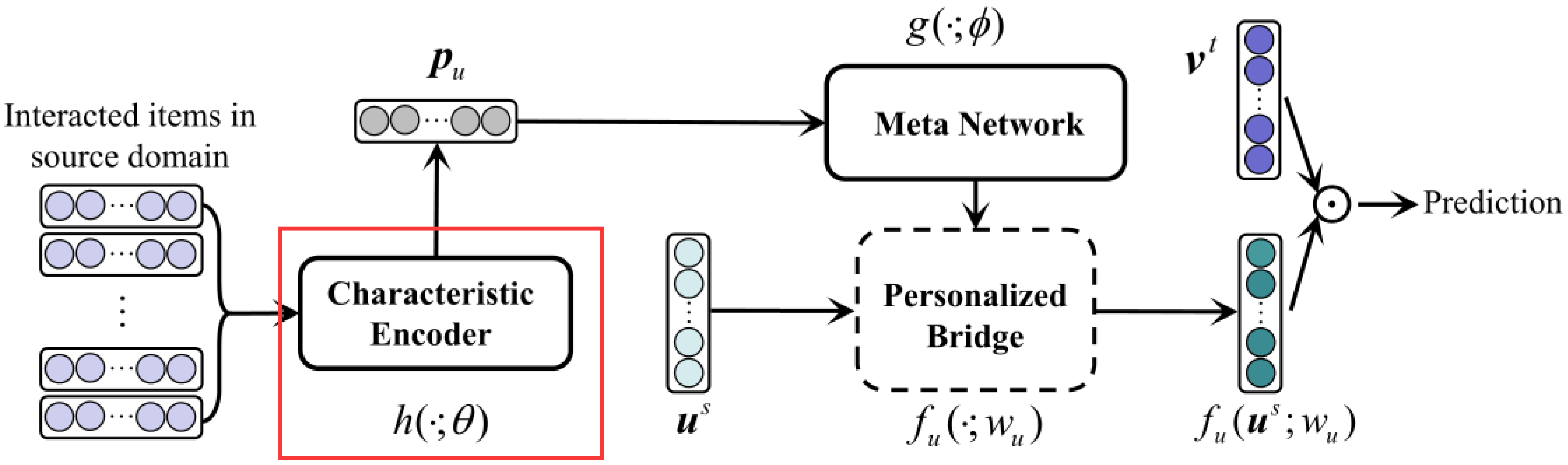


Figure 2: Personalized Transfer of User Preferences for Cross-domain Recommendation (PTUPCDR) utilizes a meta network with users' characteristic embeddings in the source domain as input to generate personalized bridge functions for each user. Then, with the personalized bridge function, we can obtain the transformed user's embeddings as the initial embeddings.

$$\mathcal{U} = \{u_1, u_2, \ldots\}$$

$$\mathcal{V} = \{v_1, v_2, \ldots\}$$

rating matrix $\mathcal{R}$.

$$\mathcal{U}^o = \mathcal{U}^s \cap \mathcal{U}^t$$

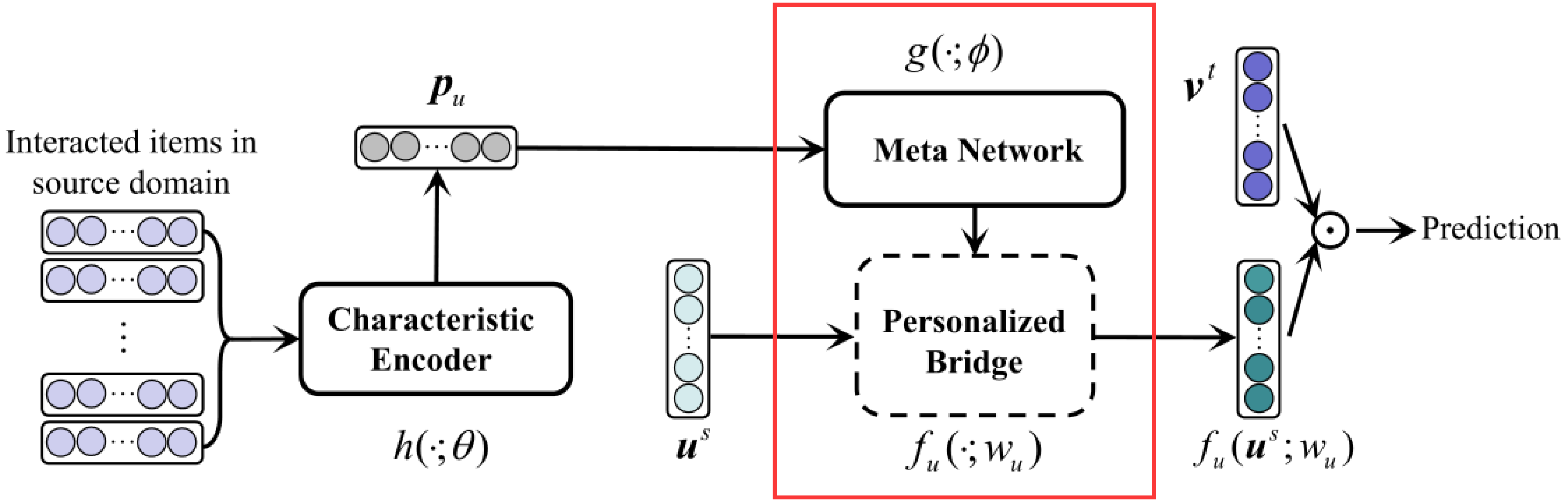
$$\mathcal{S}_{u_i} = \{v^s_{t_1}, v^s_{t_2}, \cdots, v^s_{t_n}\},$$

$$\boldsymbol{p}_{u_i} = \sum_{v^s_j \in \mathcal{S}_{u_i}} a_j \boldsymbol{v}^s_j, \tag{1}$$

$$a'_j = h(\boldsymbol{v}_j; \theta),$$
$$a_j = \frac{\exp(a'_j)}{\sum_{v^s_l \in \mathcal{S}_{u_i}} \exp(a'_l)}, \tag{2}$$

$$\min_{\boldsymbol{u},\boldsymbol{v}} \frac{1}{|\mathcal{R}|} \sum_{r_{ij} \in \mathcal{R}} (r_{ij} - \boldsymbol{u}_i \boldsymbol{v}_j)^2, \tag{8}$$

Approach



$$\mathbf{w}_{u_i} = g(\mathbf{p}_{u_i};\phi), \tag{3}$$

$$f_{u_i}(\cdot;\mathbf{w}_{u_i}), \tag{4}$$

$$\hat{\mathbf{u}}_i^t = f_{u_i}(\mathbf{u}_i^s;\mathbf{w}_{u_i}), \tag{5}$$

$$\mathcal{L} = \sum_{u_i \in \mathcal{U}^o} ||\hat{\mathbf{u}}_i^t - \mathbf{u}_i^t||^2, \tag{6}$$

$$\min_{\theta,\phi} \frac{1}{|\mathcal{R}_o^t|} \sum_{r_{ij} \in \mathcal{R}_o^t} (r_{ij} - f_{u_i}(\mathbf{u}_i^s;\mathbf{w}_{u_i})\mathbf{v}_j)^2, \tag{7}$$

Experiments

Table 1: Statistics of the cross-domain tasks (*Overlap* denotes the number of overlapping users).

CDR Tasks	Domain		Item		User			Rating	
	Source	Target	Source	Target	Overlap	Source	Target	Source	Target
Task1	Movie	Music	50,052	64,443	18,031	123,960	75,258	1,697,533	1,097,592
Task2	Book	Movie	367,982	50,052	37,388	603,668	123,960	8,898,041	1,697,533
Task3	Book	Music	367,982	64,443	16,738	603,668	75,258	8,898,041	1,097,592

Experiments

Table 2: Cold-start results (MAE and RMSE) of 3 cross-domain tasks. We report the mean results over five runs. Best results are in boldface. $*$ indicates 0.05 level, paired t-test of PTUPCDR vs. the best baselines. *Improve* **denotes relative improvement over the best baseline.**

	β	*Metric*	TGT	CMF	DCDCSR	SSCDR	EMCDR	PTUPCDR	Improve
Task1	*20%*	*MAE*	4.4803	1.5209	1.4918	1.3017	1.2350	**1.1504***	6.86%
		RMSE	5.1580	2.0158	1.9210	1.6579	1.5515	**1.5195**	2.06%
	50%	*MAE*	4.4989	1.6893	1.8144	1.3762	1.3277	**1.2804***	3.57%
		RMSE	5.1736	2.2271	2.3439	1.7477	1.6644	**1.6380**	1.59%
	80%	*MAE*	4.5020	2.4186	2.7194	1.5046	1.5008	**1.4049***	6.39%
		RMSE	5.1891	3.0936	3.3065	1.9229	1.8771	**1.8234***	2.86%
Task2	*20%*	*MAE*	4.1831	1.3632	1.3971	1.2390	1.1162	**0.9970***	10.68%
		RMSE	4.7536	1.7918	1.7346	1.6526	1.4120	**1.3317***	5.69%
	50%	*MAE*	4.2288	1.5813	1.6731	1.2137	1.1832	**1.0894***	7.93%
		RMSE	4.7920	2.0886	2.0551	1.5602	1.4981	**1.4395***	3.91%
	80%	*MAE*	4.2123	2.1577	2.3618	1.3172	1.3156	**1.1999***	8.80%
		RMSE	4.8149	2.6777	2.7702	1.7024	1.6433	**1.5916***	3.15%
Task3	*20%*	*MAE*	4.4873	1.8284	1.8411	1.5414	1.3524	**1.2286***	9.15%
		RMSE	5.1672	2.3829	2.2955	1.9283	1.6737	**1.6085***	3.90%
	50%	*MAE*	4.5073	2.1282	2.1736	1.4739	1.4723	**1.3764***	6.51%
		RMSE	5.1727	2.7275	2.6771	1.8441	1.8000	**1.7447***	3.07%
	80%	*MAE*	4.5204	3.0130	3.1405	1.6414	1.7191	**1.5784***	3.84%
		RMSE	5.2308	3.6948	3.5842	2.1403	2.1119	**2.0510***	2.88%

Experiments

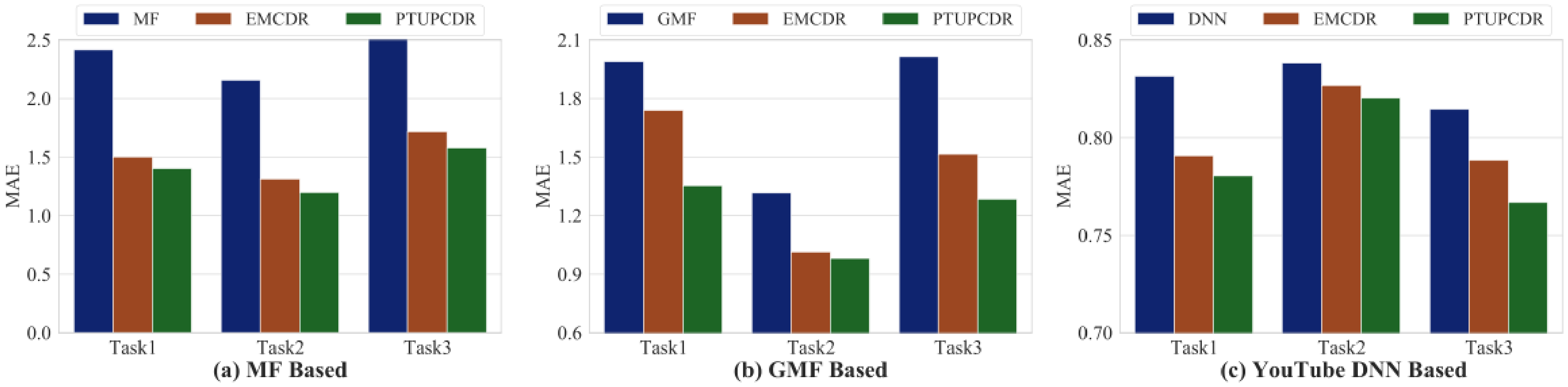


Figure 3: Generalization experiments: applying EMCDR and PTUPCDR upon three base models (a) MF, (b) GMF, and (c) YouTube DNN, and show the averaged results over five runs.

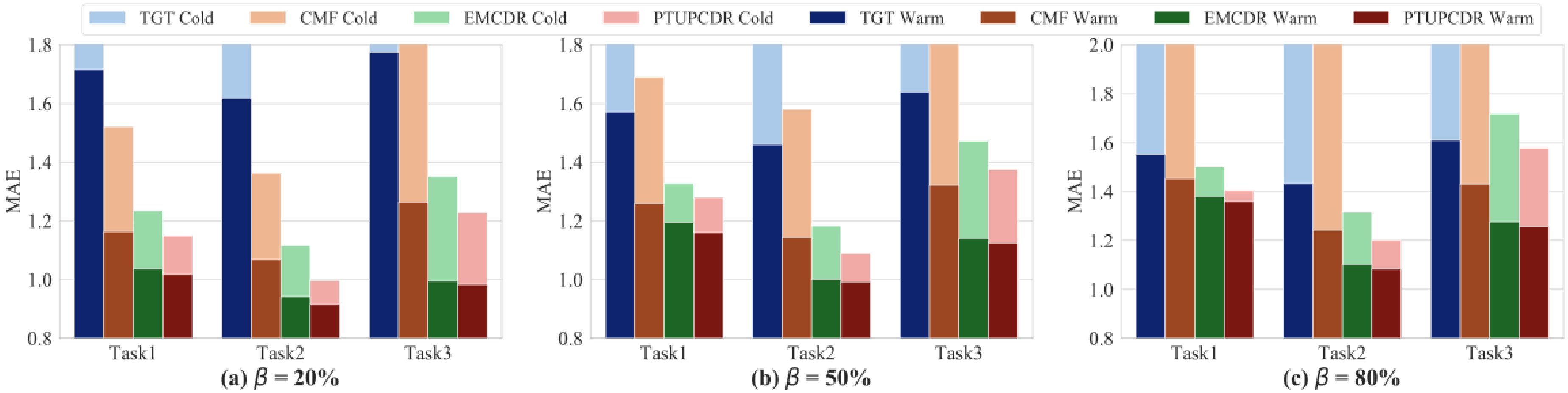


Figure 4: Warm-start experiments on TGT, CMF, EMCDR, and PTUPCDR for different proportions of test (cold-start) users β: (a) $\beta = 20\%$, (b) $\beta = 50\%$, and (c) $\beta = 80\%$. The light-colored histograms represent the performance of extreme cold-start scenario, while the dark-colored histograms represent the warm-start scenario.

Experiment

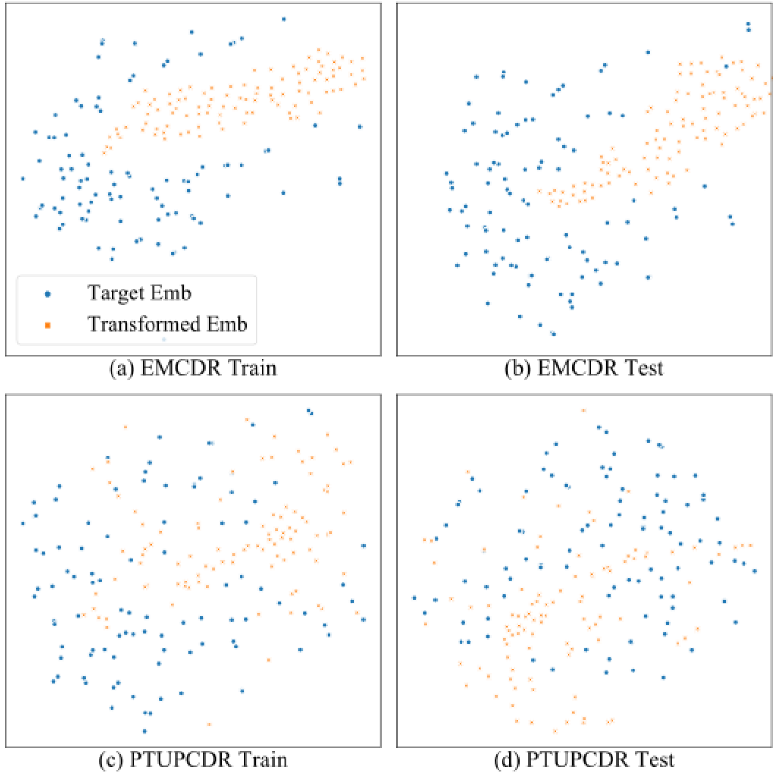


Figure 5: t-SNE visualization of randomly sampled user embeddings in target-domain feature space and transformed user embeddings. (a) and (b), (c) and (d) denotes the visualization results of EMCDR and PTUPCDR, respectively.

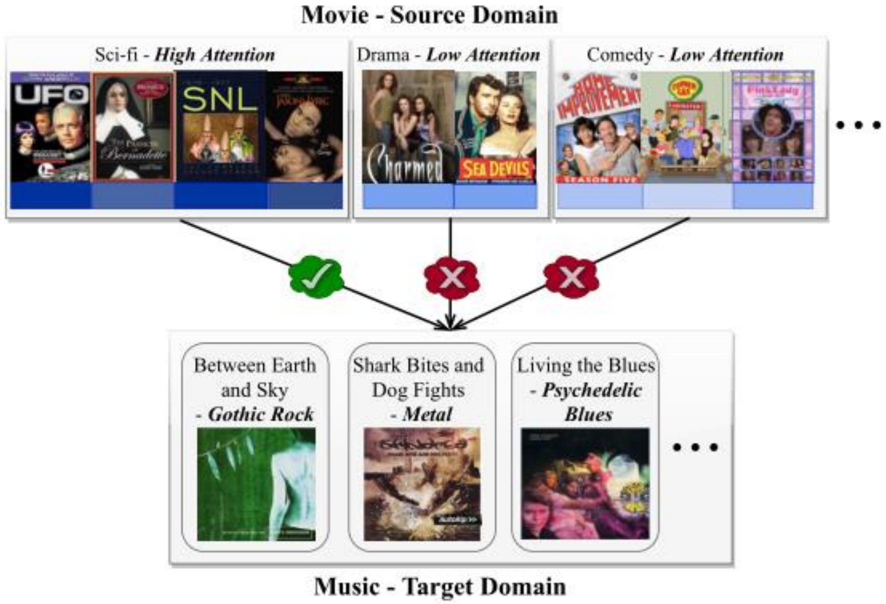


Figure 6: The color block below each movie represents the attention score. High attention items dominate the recommendation while others have little influence on results.

Thank you!