Transfer-Meta Framework for Cross-domain Recommendation to Cold-Start Users

Yongchun Zhu[1,2,3], Kaikai Ge[3], Fuzhen Zhuang[4,5], Ruobing Xie[3], Dongbo Xi[1,2],

Xu Zhang[3], Leyu Lin[3], Qing He[1,2]

1. Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS) , Institute of Computing Technology, CAS, Beijing, China

2. University of Chinese Academy of Sciences, Beijing, China

3. WeChat Search Application Department, Tencent, China

4. Institute of Artificial Intelligence, Beihang University, Beijing 100191, China

5. Xiamen Data Intelligence Academy of ICT, CAS, China

INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES

WeChat

Motivation

Cold-start Problem

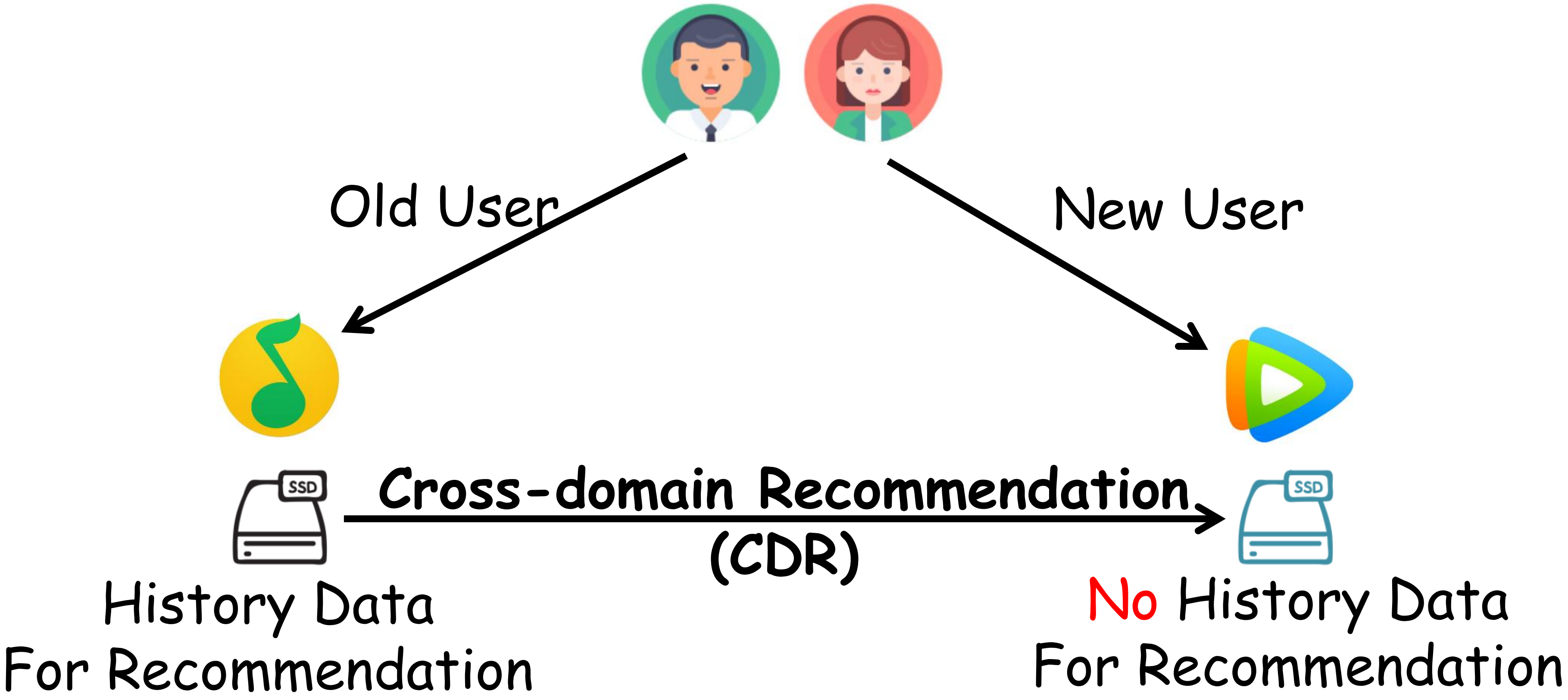
New User

Tencent Video

No History Data For Recommendation

How to recommend to these new(cold-start) users?

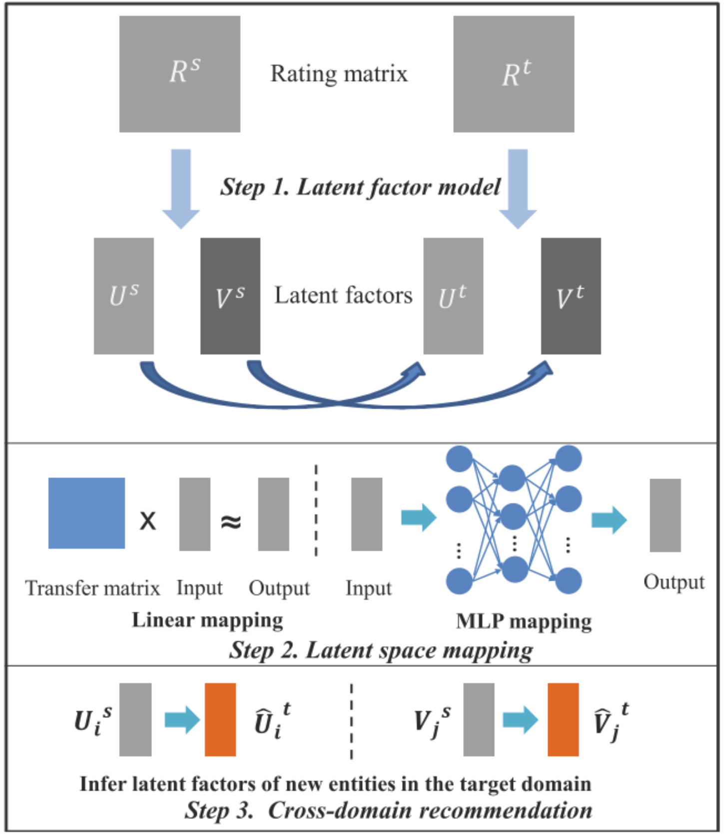
Motivation


Old User
New User
SSD
Cross-domain Recommendation
(CDR)
SSD
History Data
For Recommendation
No History Data
For Recommendation

Motivation

Unsatisfying generalization ability: EMCDR-based methods explicitly learn the mapping function by minimizing the distance between the target embedding and the mapped embedding of the overlapping users, which could be overfitting on these users.

EMCDR



Transfer-Meta Framework for CDR

Two stages

- Transfer Stage:
  - A source model
  - A target model
- Meta Stage:
  - Training a mapping function from a meta-learning perspective
  - Task-oriented loss

Algorithm 1 Transfer-Meta framework for CDR (TMCDR)

Input: Given user and item sets of source and target domains, U^s, U^t, V^s, V^t. The overlapping user set U^o. The rating matrix R^s, R^t.
Input: Task-oriented meta network f_θ.
Input: The step size (learning rate) λ, α.
Transfer Stage:

- A pre-trained source model contains $\mathrm{u}^s, \mathrm{v}^s$.
- A pre-trained target model contains $\mathrm{u}^t, \mathrm{v}^t$.

Meta Stage: utilize the source embedding of overlapping users u^s and the target item embedding v^t to optimize the task-oriented meta network f_θ.

(1) randomly initialize θ.
(2) **while** not converge **do**:
(3) sample batch of user groups $\{U_1, ..., U_n\}$ from U^o.
(4) **for** $U_i \in \{U_1, ..., U_n\}$ do:
(5) divide U_i into two disjoint sets U_a, U_b
(6) define two training sets D_a, D_b with U_a, U_b
(7) evaluate loss $\mathcal{L}_\theta$ with D_a
(8) compute updated parameter $\theta' = \theta - \lambda \frac{\partial \mathcal{L}_\theta}{\partial \theta}$
(9) evaluate loss $\mathcal{L}_{\theta'_i}$ with D_b
(10) update $\theta = \theta - \alpha \sum_{U_i \in \{U_1,...,U_n\}} \frac{\partial \mathcal{L}_{\theta'_i}}{\partial \theta}$
(11) **end while**

Test Stage: for a cold-start user u, we use $f_\theta(\mathrm{u}^s)$ as the user embedding for prediction.

Experiments

Dataset

Dataset	CDR Tasks			Item		User			Rating	
	Scenarios	Source	Target	Source	Target	Overlap	Source	Target	Source	Target
Amazon	Scenario 1(S1)	apps	video	13,209	10,672	894	87,271	24,303	752,937	231,779
	Scenario 2(S2)	home	tools	16,638	6,038	6,038	66,519	16,638	551,682	134,475
	Scenario 3(S3)	movies	cds	50,052	64,443	18,031	123,960	75,258	1,697,533	1,097,591
	Scenario 4(S4)	books	movies	98,700	17,798	37,388	603,668	123,960	6,466,068	1,420,441
Douban	Scenario 5(S5)	movie	music	16,281	15,581	31,360	41,751	33,173	8,113,064	2,912,666
	Scenario 6(S6)	music	book	19,921	16,999	27,779	33,281	34,166	3,064,996	1,603,329

Metrics: AUC, NDCG

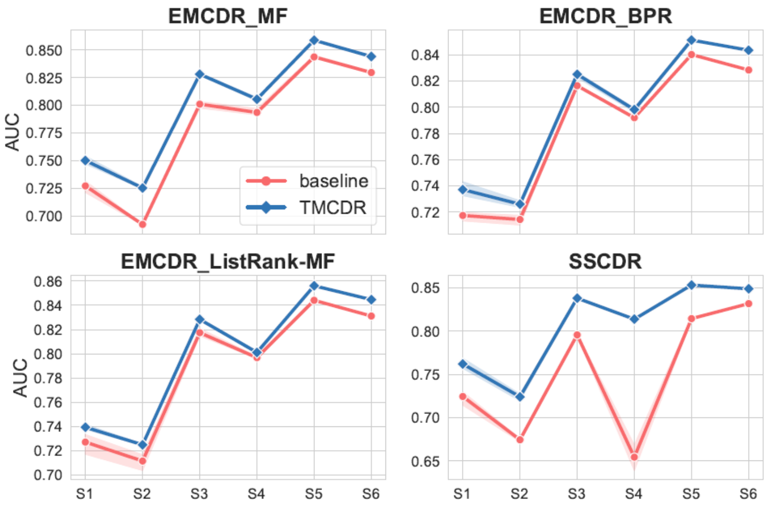
Baselines: CMF, BPR, ListRank-MF, CML, CST, EMCDR, SSCDR

Experiments

Overall Performance

Method	AUC (Scenario1)	NDCG@10 (Scenario1)	AUC (Scenario2)	NDCG@10 (Scenario2)
CMF	0.6490	0.1696	0.6996	0.2076
BPR	0.7226	0.2182	0.7160	0.2379
ListRank-MF	0.6648	0.1709	0.7232	0.2204
CML	0.6470	0.1408	0.6986	0.2147
CST	0.7240	0.2137	0.7124	0.2324
SSCDR	0.7245	0.0089	0.6745	0.0013
EMCDR_MFori	0.6942	0.1978	0.6511	0.1747
EMCDR_MF	0.7271	0.2103	0.6923	0.1985
TMCDR_MF	**0.7501***	**0.2246***	**0.7253***	**0.2427***
	Scenario3		Scenario4	
CMF	0.7769	0.3066	0.7295	0.2349
BPR	0.7737	0.3065	0.7199	0.2150
ListRank-MF	0.7640	0.2902	0.7409	0.2277
CML	0.8191	**0.3548**	0.7857	0.2647
CST	0.7995	0.2960	0.7842	0.2563
SSCDR	0.7956	0.3080	0.6545	0.1628
EMCDR_MFori	0.7273	0.2284	0.7307	0.1990
EMCDR_MF	0.8011	0.3055	0.7936	0.2670
TMCDR_MF	**0.8282***	0.3334	**0.8056***	**0.2775***
	Scenario5		Scenario6	
CMF	0.8465	0.3420	0.8339	0.3764
BPR	0.8108	0.3283	0.8138	0.3659
ListRank-MF	0.8136	0.3106	0.8191	0.3281
CML	0.8466	0.3409	0.8405	0.3707
CST	0.8524	0.3405	0.8406	0.3742
SSCDR	0.8144	0.2925	0.8317	0.3644
EMCDR_MFori	0.7307	0.1990	0.7627	0.2703
EMCDR_MF	0.8438	0.3322	0.8297	0.3702
TMCDR_MF	**0.8589***	**0.3483***	**0.8442***	**0.3778***

Generalization Experiments



Thanks Q & A

Name: Yongchun Zhu | Email: zhuyongchun18s@ict.ac.cn

中国科学院计算技术研究所
INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES

WeChat