Adversarial Preference Learning with Pairwise Comparisons

Zitai Wang[1,2], Qianqian Xu[3], Ke Ma[1,2], Yangbangyan Jiang[1,2], Xiaochun Cao[1,2], Qingming Huang[2,3,4*]

[1]Institute of Information Engineering, CAS [2]University of Chinese Academy of Sciences

[3]Institute of Computing Technology, CAS [4]BDKM, CAS

wang22ti@gmail.com, xuqianqian@ict.ac.cn, {make, jiangyangbangyan, caoxiaochun}@iie.ac.cn, qmhuang@ucas.ac.cn

Motivation

Collaborative ranking (CR) is a ranking-based variant of collaborative filtering. For example, traditional CR generative model:

$$\min_{\Theta_g} \sum_{(u,i,j)\in\mathcal{T}} \mathcal{L}_g\left(\mathcal{P}\left(\underbrace{\sigma(s_{ui}-s_{uj})}_{\text{score difference}}, \overbrace{y_{uij}}^{\text{label of pairwise comparison}} | \Theta_g\right)\right)$$

Pairwise comparisons are adopted to avoid calibration drawback (the same rating represents different user preference).

static learning paradigm with a fixed score function

→ restrict its further improvement of precision.

Framework

Learning **dynamically** against increasing difficulty and adversarial attacks:

- **Generator** learning non-linear score function:

$$g_\theta(u,i) = \sigma\left(\underbrace{g_m}_{\text{multi-FC}}\left(\underbrace{g_E^{\text{user}}(u), g_E^{\text{item}}(i)}_{\text{embedding}}\right)\right)$$

multi-FC, embedding layer

Score difference:

$$\Delta G_{\theta,t} = g_\theta(u,i) - g_\theta(u,j), \ \forall\ t=(u,i,j)\in\mathcal{T}$$

- **Discriminator** providing stricter supervision signals:

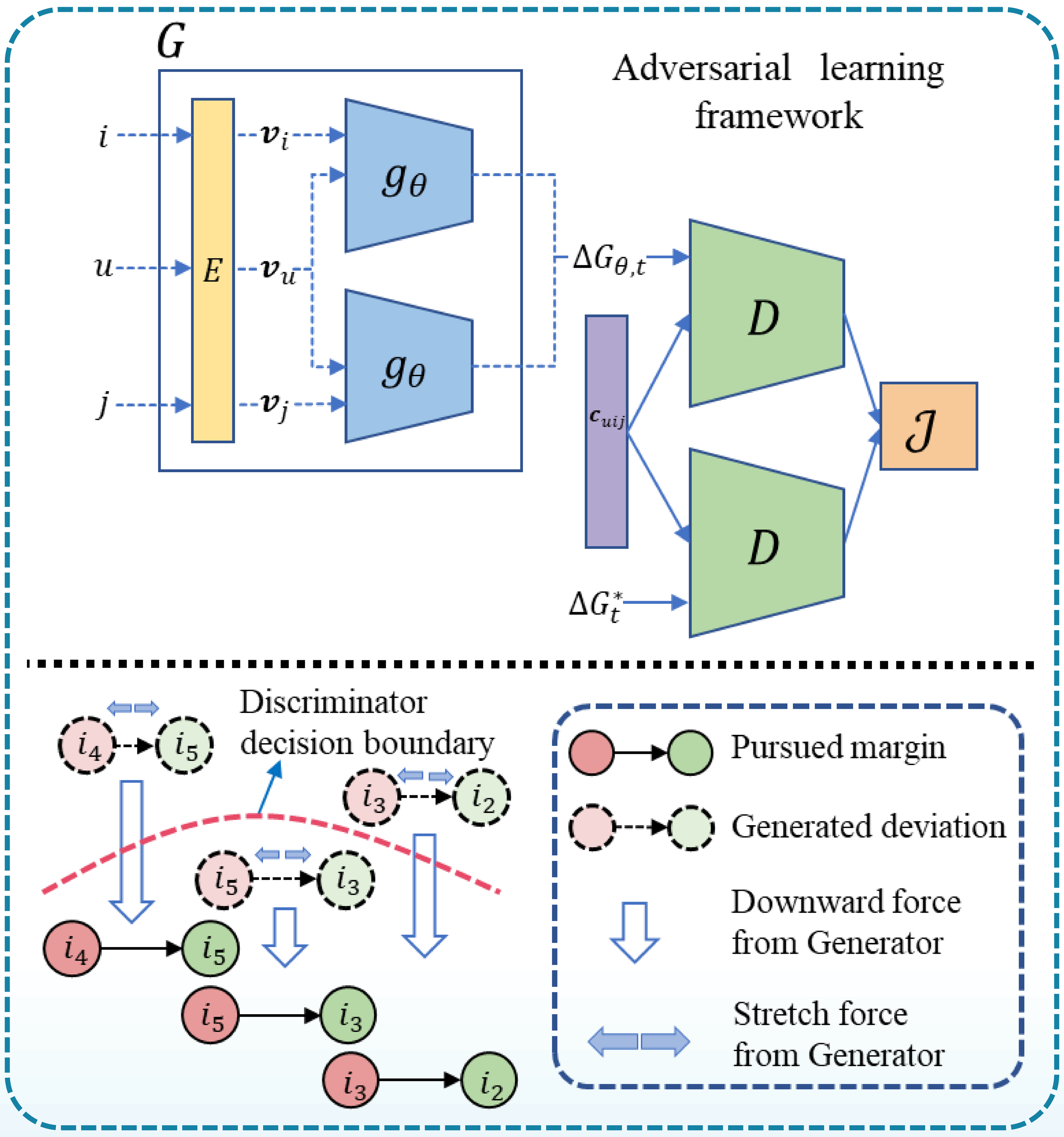
$$d_\phi\left(\Delta G_t, \underbrace{\mathbf{c}_{uij}}_{\text{code of triplet } (u,i,j)}, z_{uij}\right) = \sigma\left(\underbrace{z_{uij}}_{\text{Indicator of } \Delta G_t} \cdot f_m\left(\Delta G_t, \mathbf{c}_{uij}\right)\right)$$

where

$$\Delta G_t = \begin{cases} \Delta G_{\theta,t} & \text{if } z_{uij} = -1, \\ \underbrace{\Delta G_t^*}_{\text{ideal margin}} & \text{if } z_{uij} = \ \ 1, \end{cases}$$

- **Objective Function:**

$$\min_\theta \max_\phi J(G,d) = \mathbb{E}_{\Delta G_t^* \sim \mathcal{P}_{\Delta G^*}}\left[\log d_\phi\left(\Delta G_t^*\right)\right] + \mathbb{E}_{t\sim\mathcal{P}_\mathcal{T}}\left[\log\left(1-d_\phi\left(\Delta G_{\theta,t}\right)\right)\right]$$



Advantage

In view of probability, if we fix d and update g, the objective can be simplified by Bayes' rule:

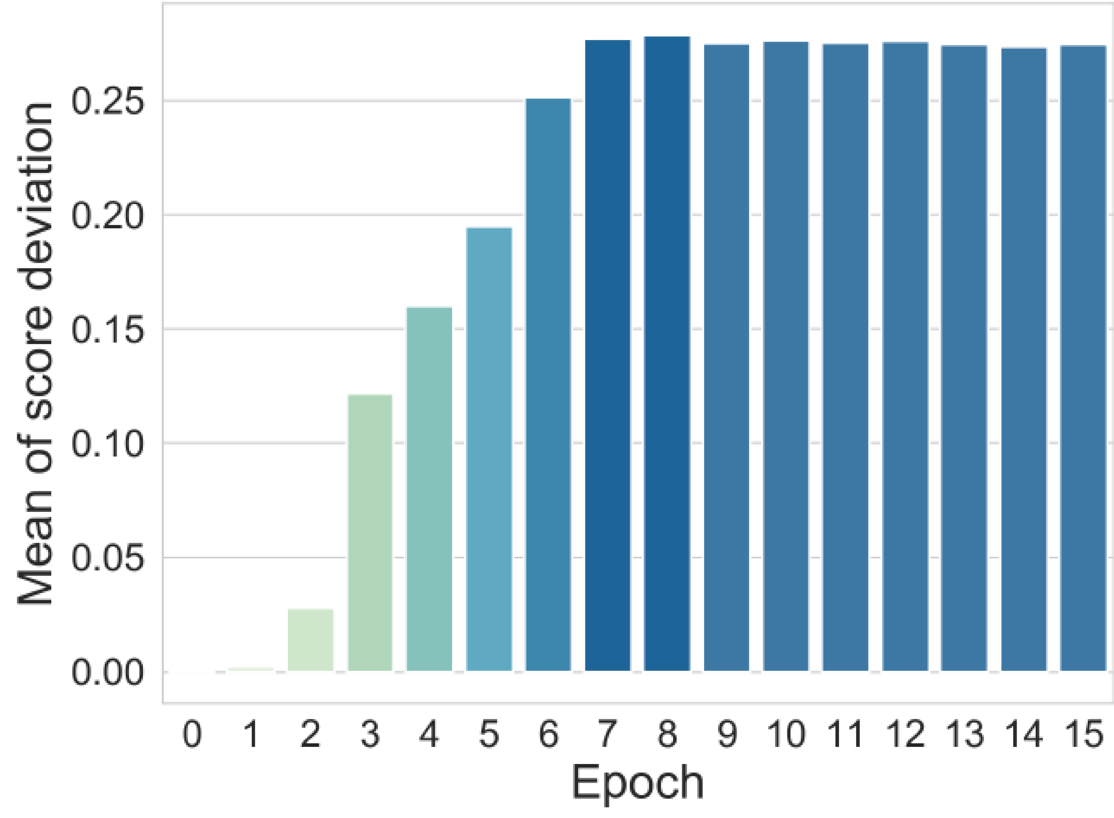
$$\min_\theta \sum_{(u,i,j)\in\mathcal{T}} \log\left(\mathcal{P}\left(\Delta G_{\theta,t}, y_{uij} | z_{uij} = -1; \phi^*\right)\right)$$

If we let $\mathcal{L}_g'(\cdot) = \log(\cdot)$ and $\sigma'(\cdot) = d_{\phi^*}(\cdot)$:

$$\min_\theta \sum_{(u,i,j)\in\mathcal{T}} \mathcal{L}_g'\left(\mathcal{P}\left(\sigma'\left(\Delta G_{\theta,t}\right), y_{uij}\right)\right)$$

Intuitively, our framework generalizes traditional generative methods, which explains its superior performance:

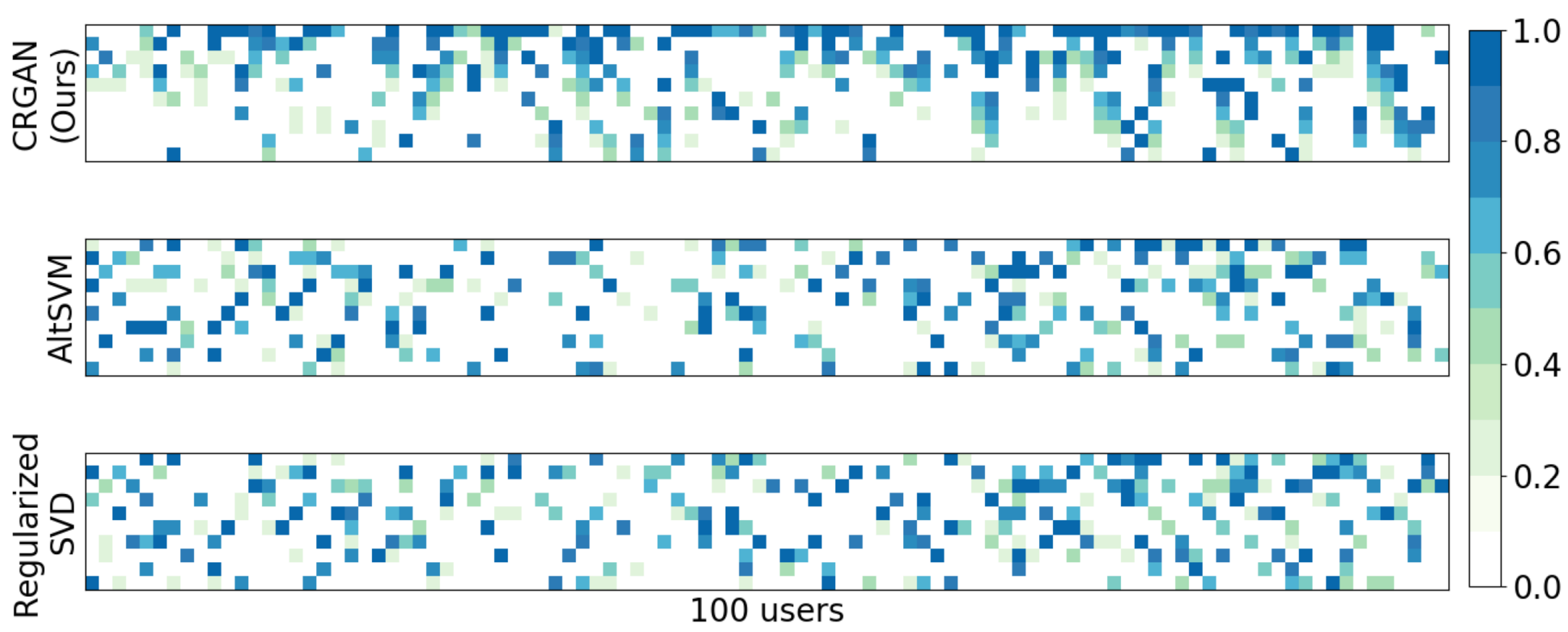
- Deep generator → adaptive score function
- Stricter supervision signals from discriminator
  → score differences continually approximate the ideal margin:



Experiment

	HR@10↑	P@10↑	NDCG@10↑	AUC@10↑	MAP@10↑	MRR@10↑
IRGAN	0.8511	0.2642	0.2670	0.5989	0.2858	0.1020
MLP	0.9598	0.3519	0.3909	0.7271	0.3938	0.1491
GMF	0.9175	0.3437	0.3580	0.7287	0.3762	0.1331
NeuMF	0.9416	0.3590	0.3757	0.7340	0.3858	0.1382
CoFiRank	0.9256	0.3354	0.3598	0.7152	0.3717	0.1357
RegularizedSVD	0.8934	0.3404	0.3505	0.7287	0.3702	0.1298
RankBasedSVD	0.9256	0.3318	0.3510	0.7042	0.3639	0.1318
LCR	0.9437	0.3503	0.3744	0.7245	0.3830	0.1404
Primal-CR	0.9577	0.3899	0.4205	0.7160	0.4128	0.1539
Primal-CR++	0.9575	0.3896	0.4199	0.7165	0.4139	0.1526
AltSVM	0.9618	0.3899	0.4209	0.7186	0.4136	0.1542
Global Ranking	0.9678	0.3795	0.4045	0.7353	0.4083	0.1483
CRGAN(ours)	0.9839	0.4678	0.4864	0.7559	0.4753	0.1688

Results on MovieLens100K



Ranking results for the top-10 test items on Netflix. Grids with deeper color represents higher prediction score. Obviously, our method obtains more grids with deep colors, and they tend to hit the top half of the figure.

P2B-08