Leave No One Behind: Fairness-Aware Cross-Domain Recommender Systems for Non-Overlapping Users
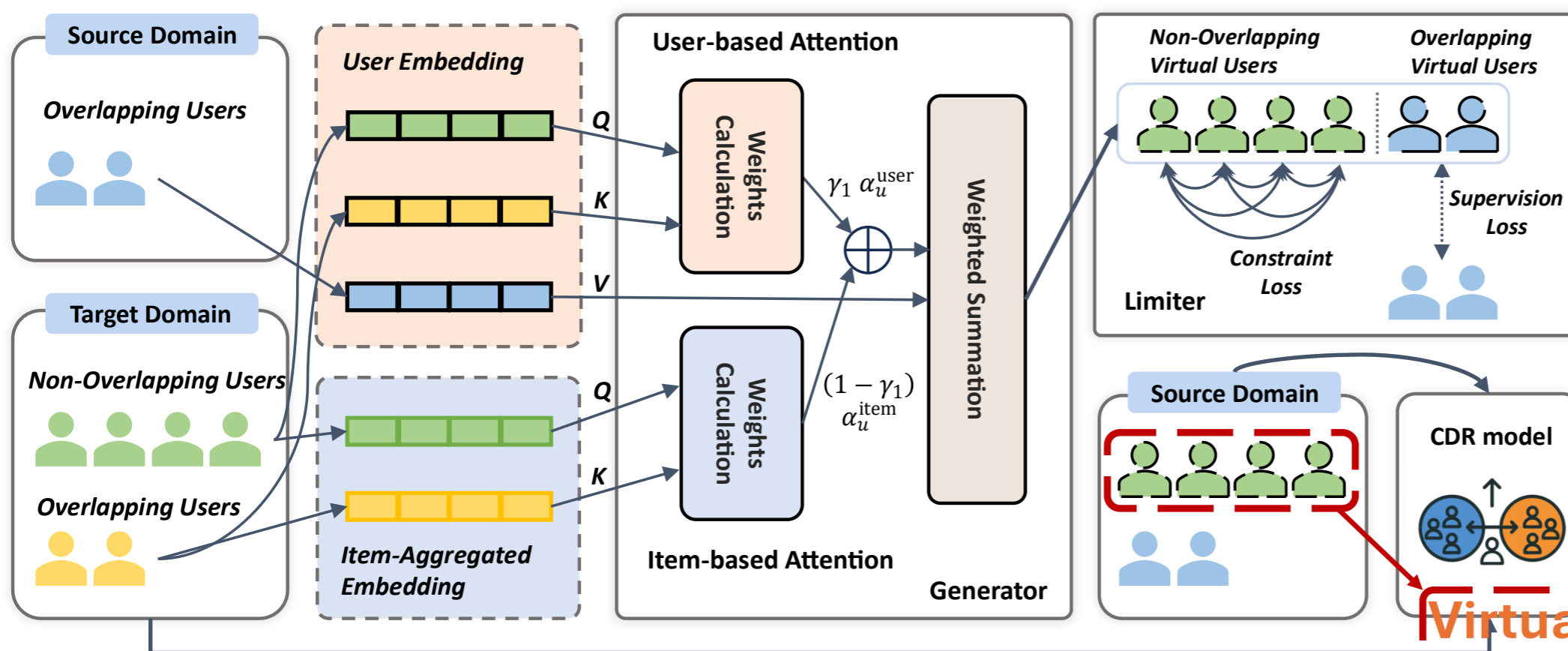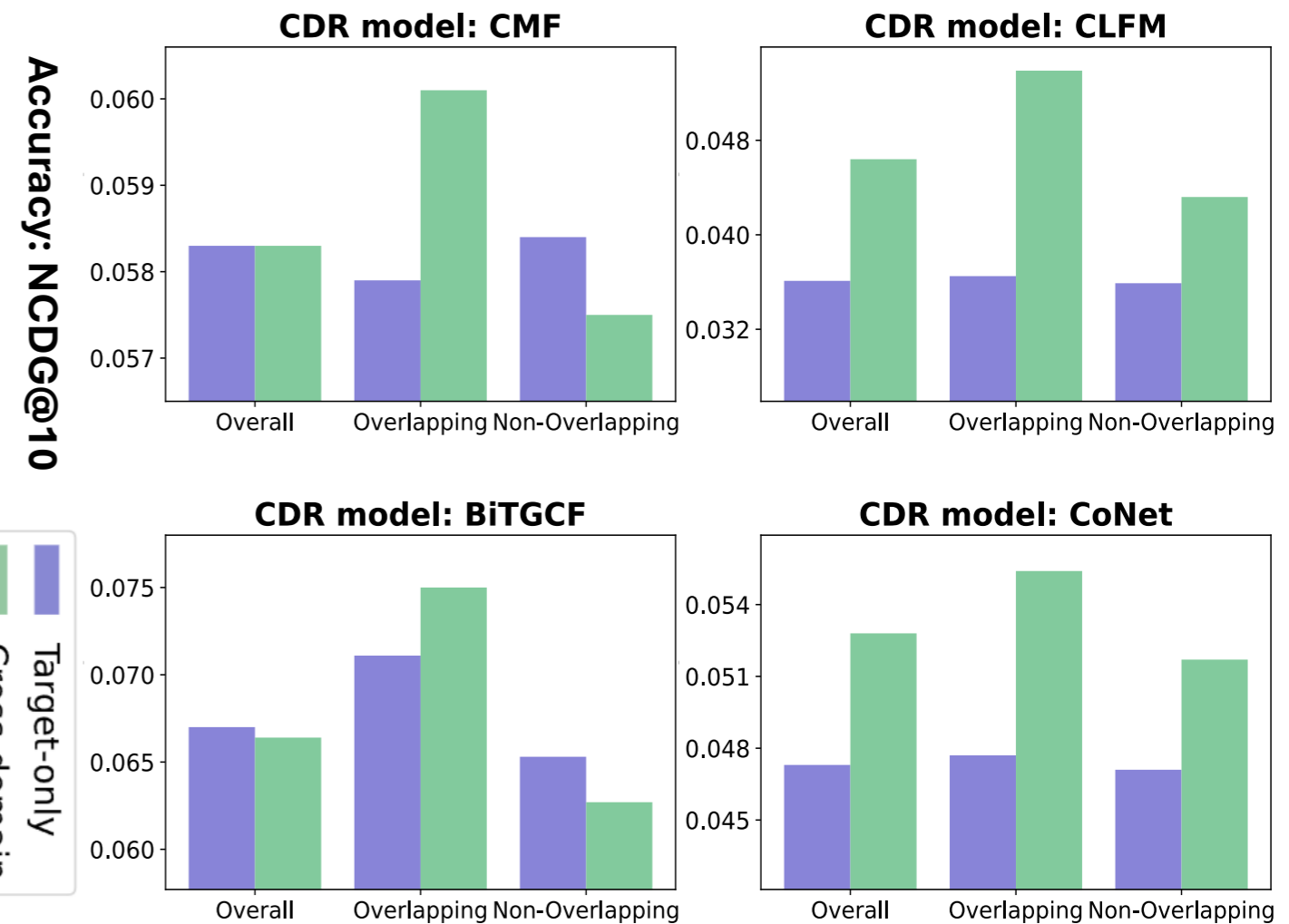
Weixin Chen
Supervisor: Prof. CHEN, Li
Department of Computer Science
Hong Kong Baptist University

香港浸會大學
HONG KONG BAPTIST UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE
計算機科學系

Research Problem:
Unfairness in Cross-Domain Recommendation

- **Cross-domain recommendation (CDR)** leverages **overlapping users** to transfer knowledge from the source domain and improve target-domain performance.

- **CDR overlapping bias:** **non-overlapping users** see less improvement or even performance degradation compared to overlapping users.

This unfairness may erode user trust, and, consequently, negatively impact business engagement and revenue.

Accuracy: NDCG@10

CDR model: CMF · CDR model: CLFM · CDR model: BiTGCF · CDR model: CoNet

Target-only / Cross-domain

Overall · Overlapping · Non-Overlapping

Methodology

Source Domain — Overlapping Users
Target Domain — Non-Overlapping Users, Overlapping Users
User Embedding · Item-Aggregated Embedding
User-based Attention · Item-based Attention
Weights Calculation
Weighted Summation
Generator
Non-Overlapping Virtual Users · Overlapping Virtual Users
Constraint Loss · Supervision Loss · Limiter
Source Domain · CDR model

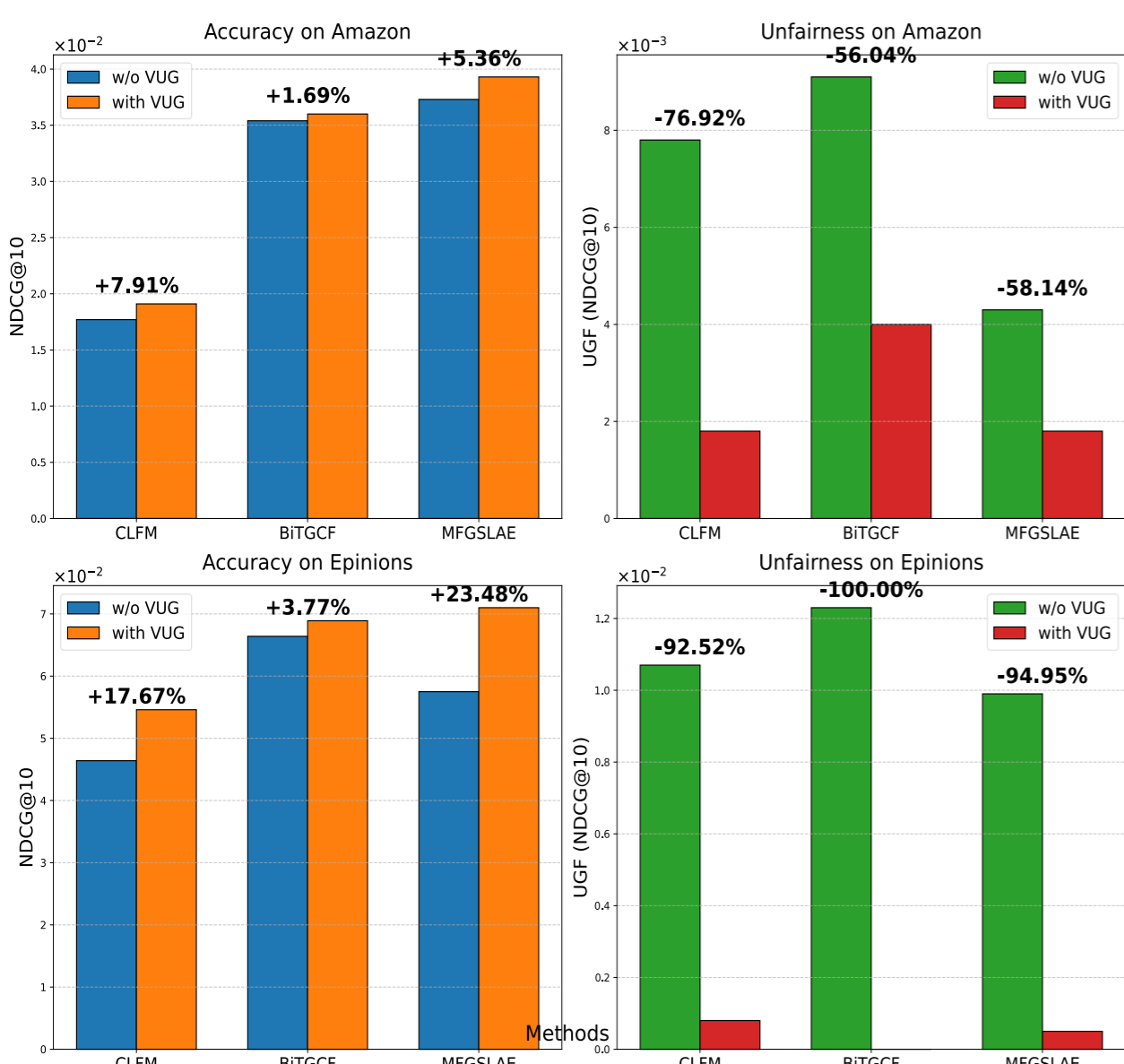$\gamma_1 \, \alpha_u^{\text{user}}$, $(1-\gamma_1)\,\alpha_u^{\text{item}}$

Key idea: generate the source embeddings for **non-overlapping users** in target domain, so they can also enjoy CDR learning:

$$\mathbf{e}_{\text{non}}^{S'} = \sum_{u \in \mathcal{U}^o} \alpha_u \left(\mathbf{W}_v \mathbf{e}_u^{S} + \mathbf{b}_v \right)$$

Virtual source embeddings for non-overlapping users in target domain

Experiments

RQ1: Overall Performance

Accuracy on Amazon: CLFM +7.91%, BiTGCF +1.69%, MFGSLAE +5.36%
Unfairness on Amazon: CLFM −76.92%, BiTGCF −56.04%, MFGSLAE −58.14%
Accuracy on Epinions: CLFM +17.67%, BiTGCF +3.77%, MFGSLAE +23.48%
Unfairness on Epinions: CLFM −92.52%, BiTGCF −100.00%, MFGSLAE −94.95%
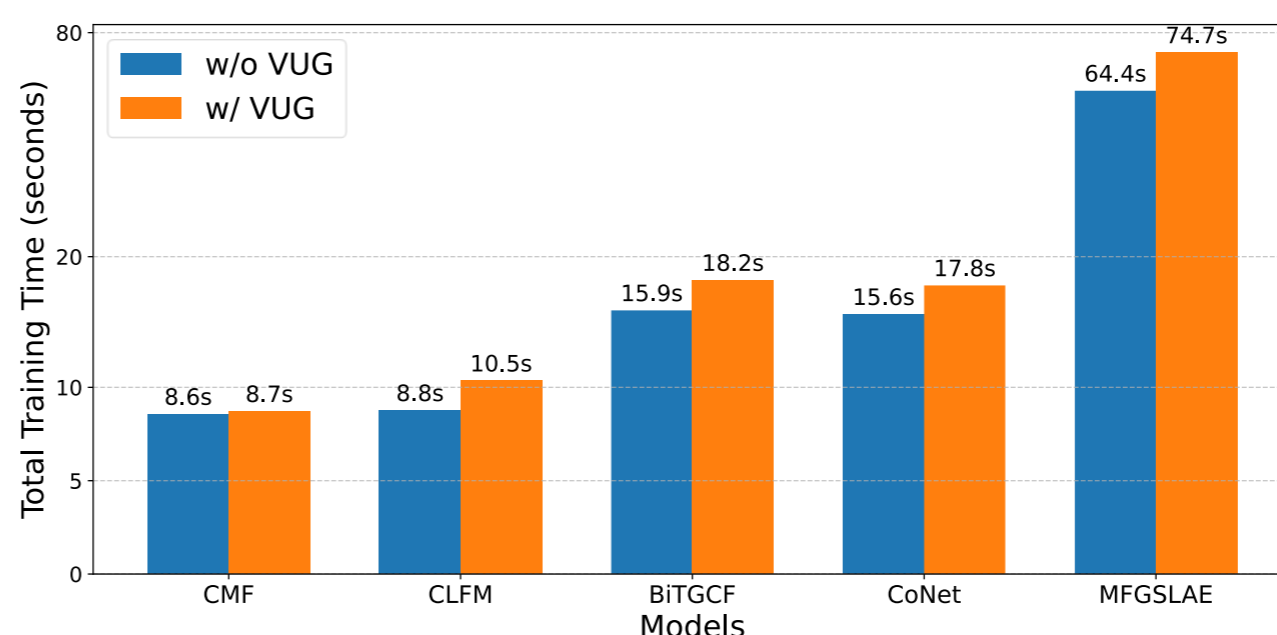
w/o VUG · with VUG

UGF measures **absolute accuracy difference** (*e.g.,* NCDG@10) between overlapping and non-overlapping users.

RQ2: Ablation Study

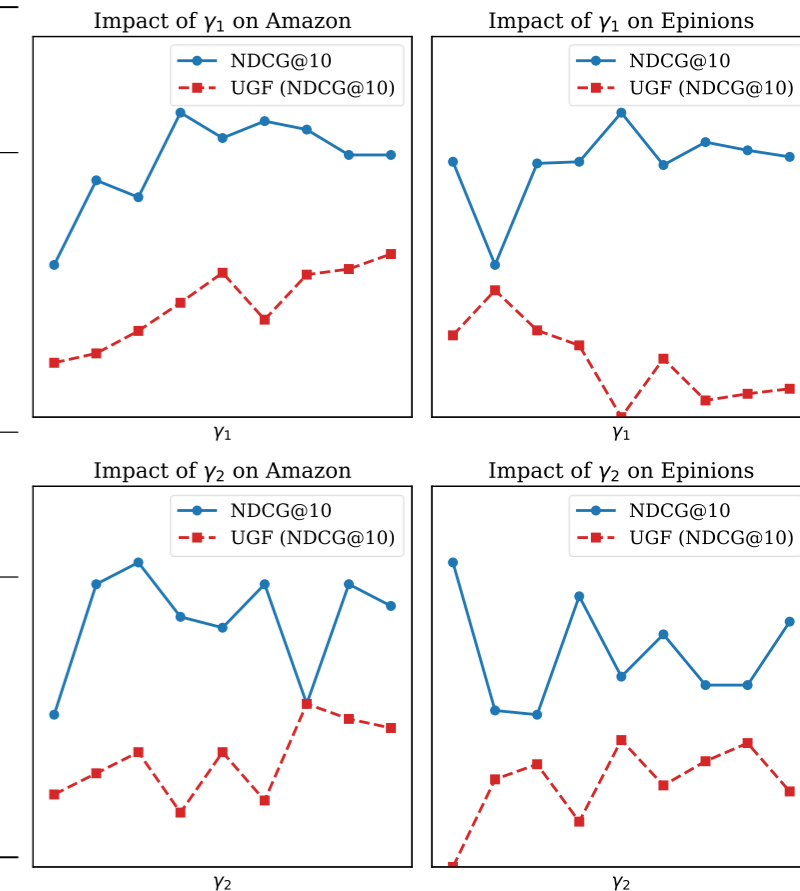| Method	Accuracy (*larger is better*)			
	HR@10	HR@20	NDCG@10	NDCG@20
w/o $\mathcal{L}_{constrain}$	0.1302	0.1974	0.0655	0.0821
w/o $\mathcal{L}_{super}$	0.1355	0.1985	0.0707	0.0864
w/o α_u^{user}	0.1298	0.1974	0.0662	0.0828
w/o α_u^{item}	0.1287	0.1963	0.0658	0.0824
VUG	**0.1363**	**0.2001**	**0.0710**	**0.0865**

Method	UGF (*smaller is better*)			
	HR@10	HR@20	NDCG@10	NDCG@20
w/o $\mathcal{L}_{constrain}$	0.0010	0.0168	0.0005	0.0033
w/o $\mathcal{L}_{super}$	0.0281	0.0152	0.0129	0.0096
w/o α_u^{user}	0.0020	0.0095	0.0015	0.0034
w/o α_u^{item}	0.0041	0.0112	**0.0005**	0.0035
VUG	**0.0003**	**0.0015**	**0.0005**	**0.0014**

RQ4: Efficiency

Total Training Time (seconds): w/o VUG · w/ VUG
CMF 8.6s / 8.7s · CLFM 8.8s / 10.5s · BiTGCF 15.9s / 18.2s · CoNet 15.6s / 17.8s · MFGSLAE 64.4s / 74.7s

RQ3: Parameter Sensitivity

Impact of γ_1 on Amazon · Impact of γ_1 on Epinions
Impact of γ_2 on Amazon · Impact of γ_2 on Epinions
NDCG@10 · UGF (NDCG@10)

Homepage: