Tiger: Transferable Interest Graph Embedding for Domain-Level Zero-Shot Recommendation

Advisor :Jia-Ling, Koh

Speaker : Shu-Ming, Yu

Source : CIKM’22

Date : 2022/12/20

Outline

- Introduction
- Method
- Experiment
- Conclusion

Outline

- Introduction
  - Domain-Level Zero-Shot
  - Knowledge Graph
  - GCN
- Method
- Experiment
- Conclusion

Domain-level Zero-shot

- It’s a cold-start problem.
- It’s may be caused by:
    - The domain being newly launched without existing user-item interactions

      Ex: A news provider wants to insert some personalized advertisements.
    - Users’ behaviors being too sensitive to collect for training

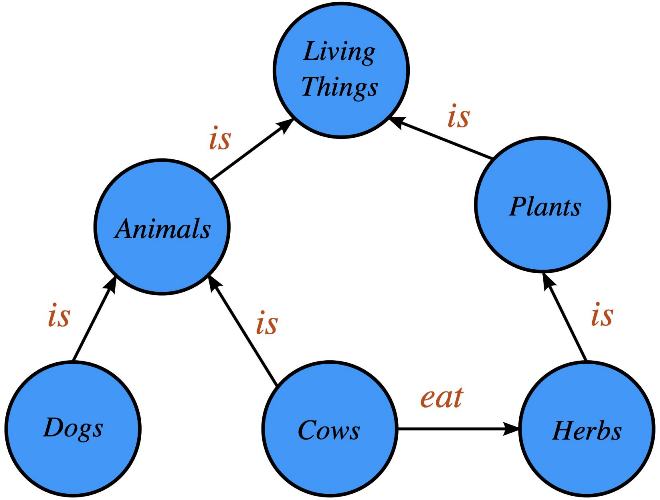
      Ex: The data has user’s email history or password

Knowledge graph

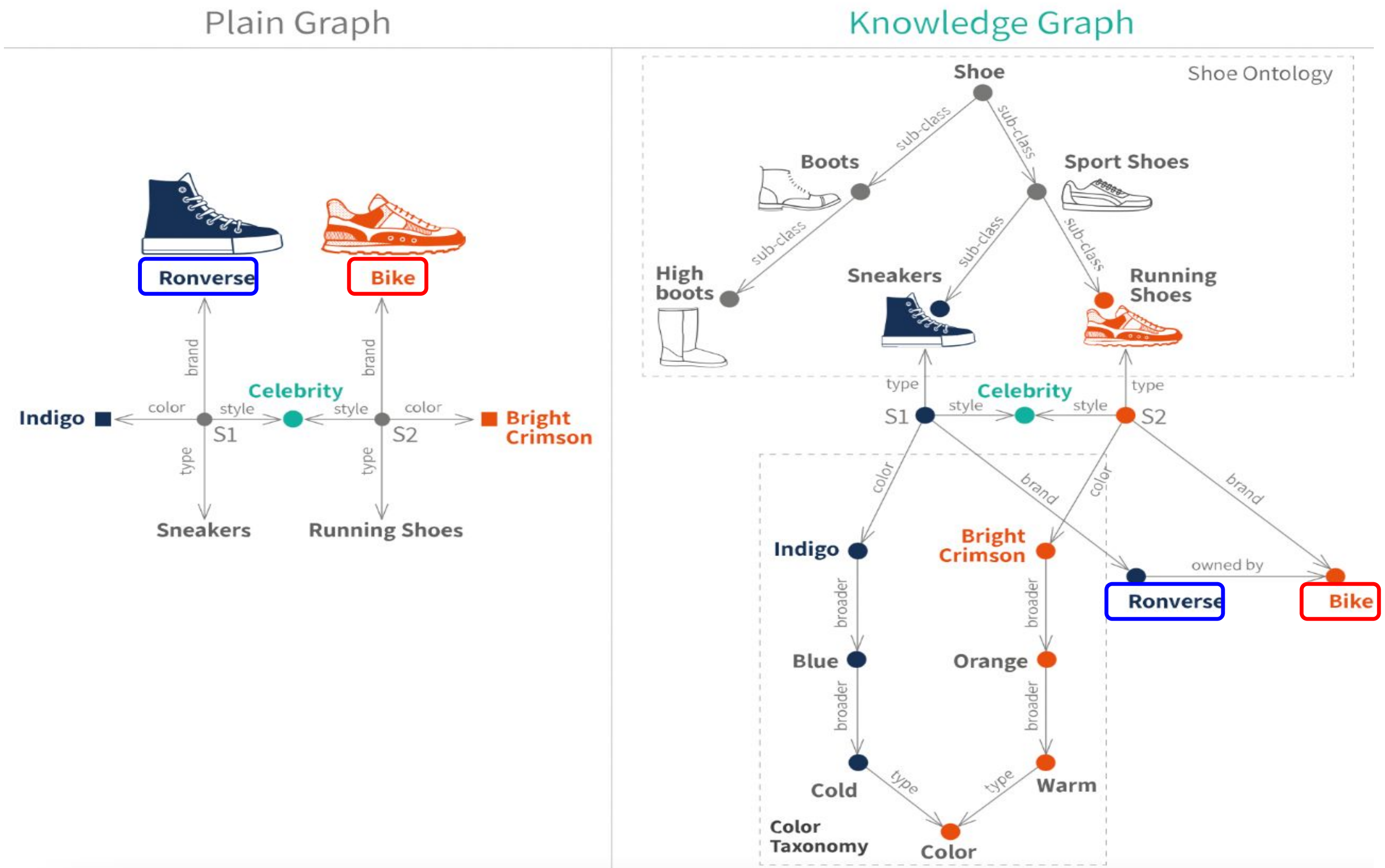
- A semantic network which represent the collections of related entities.
- It uses a graph-structured data model.
- Usually form as entity-relation-entity.

$$\mathcal{G} = \{(h, r, t) | h, r \in \mathcal{E}, r \in \mathcal{R}\}$$

h: head entity, r: relation, t: tail entity

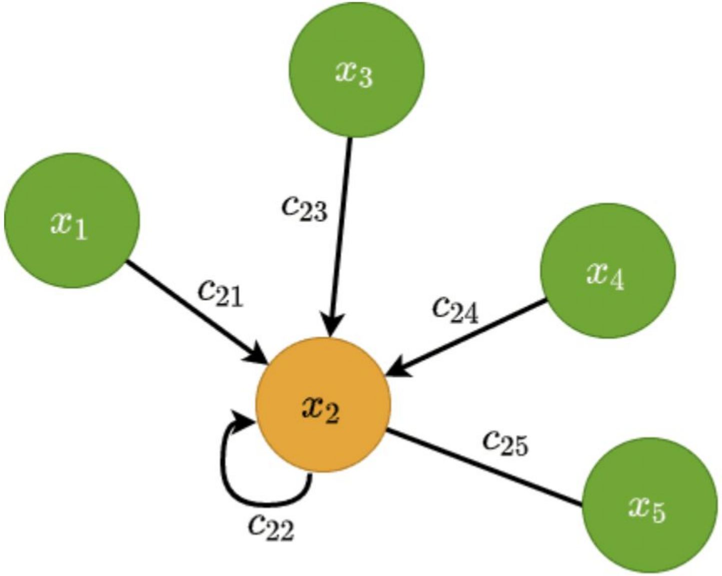


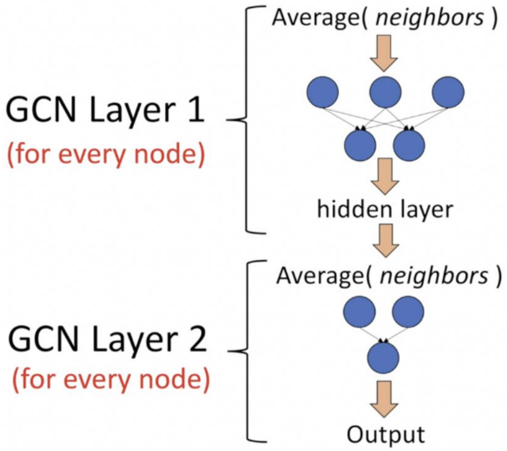
Knowledge graph



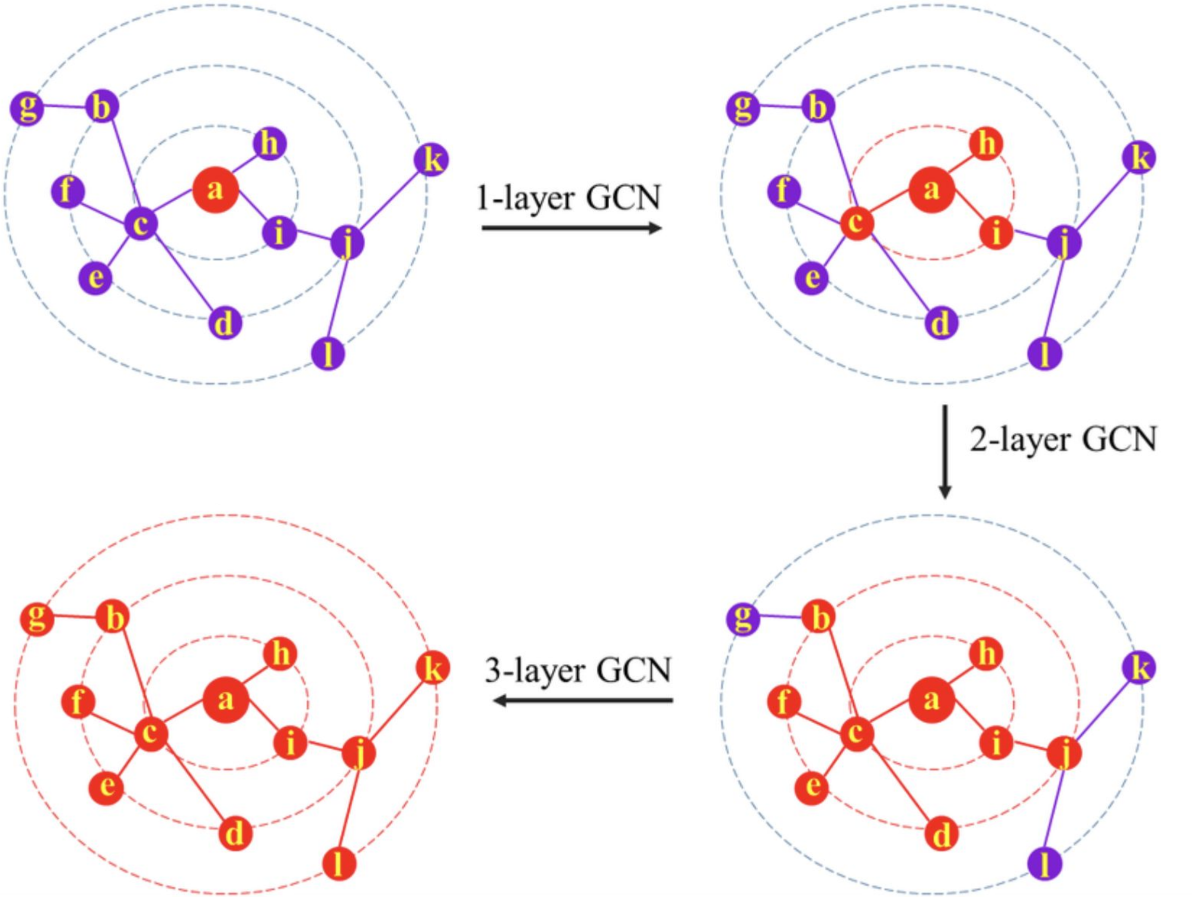
GCN

- For each node, the feature information is from all its neighbors and itself.
- The output of a layer will be treated as the input for the next layer





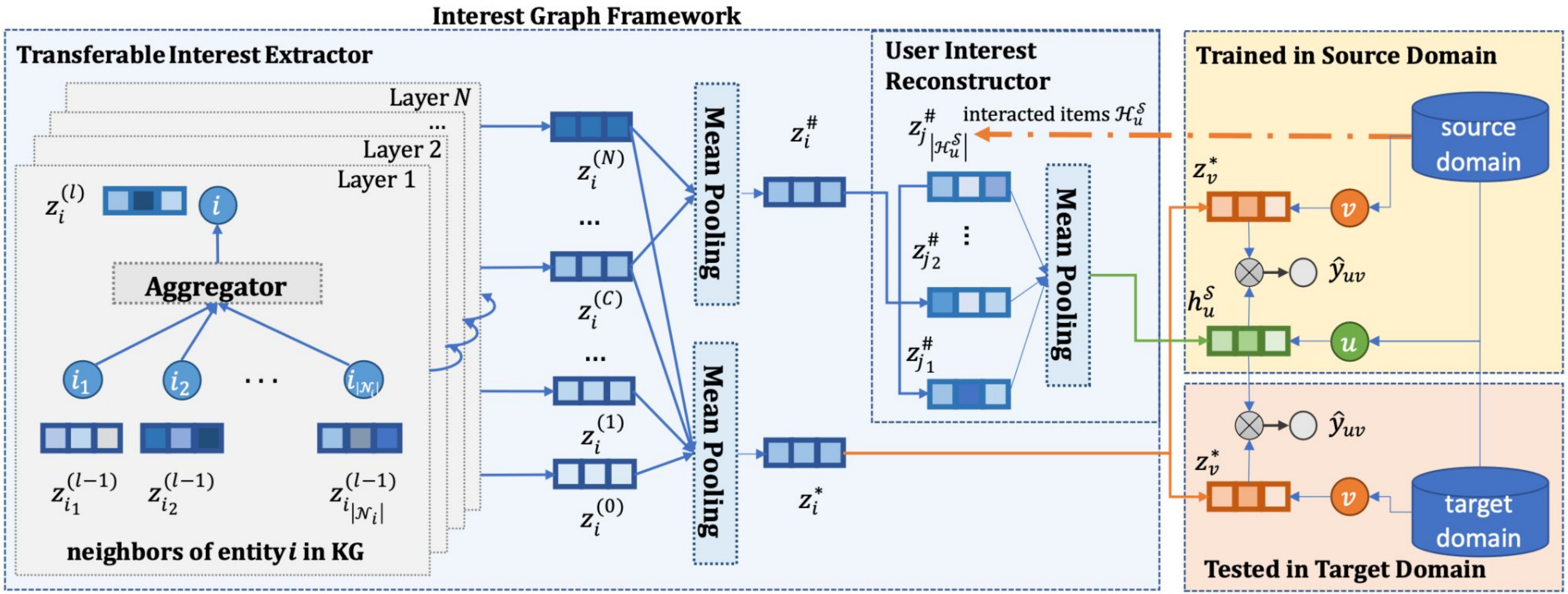
GCN



Outline

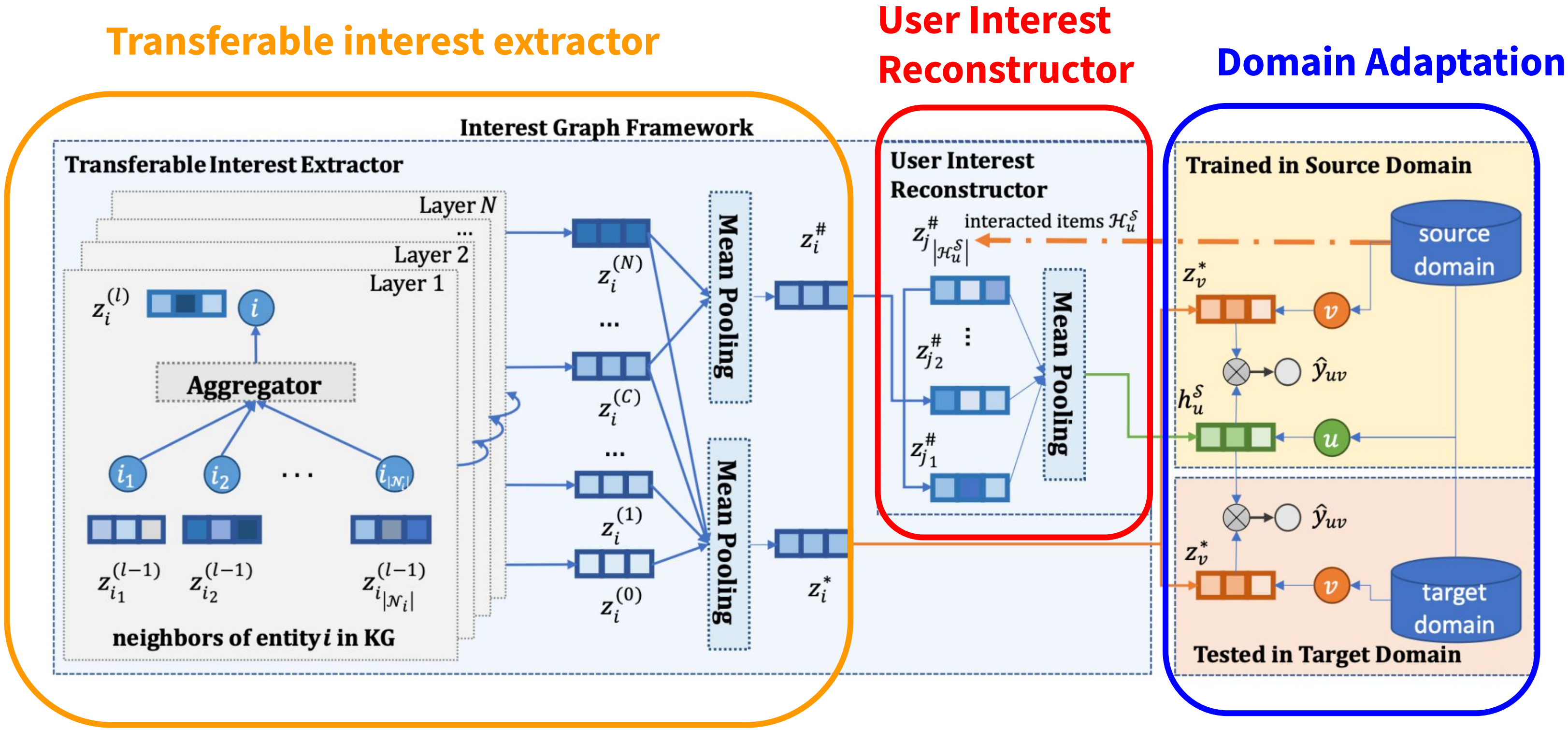
- Introduction
- Method
  - Transferable Interest Extractor
  - User Interest Reconstructor
  - Domain Adaptation
- Experiment
- Conclusion

Method



Interest Graph Framework
Transferable Interest Extractor
Layer N
Layer 2
Layer 1
$z_i^{(l)}$
Aggregator
neighbors of entity i in KG
$z_i^{(N)}$
$z_i^{(C)}$
$z_i^{(1)}$
$z_i^{(0)}$
Mean Pooling
Mean Pooling
$z_i^{\#}$
z_i^{*}
User Interest Reconstructor
interacted items $\mathcal{H}_u^S$
Mean Pooling
Trained in Source Domain
source domain
h_u^S
$\hat{y}_{uv}$
Tested in Target Domain
target domain

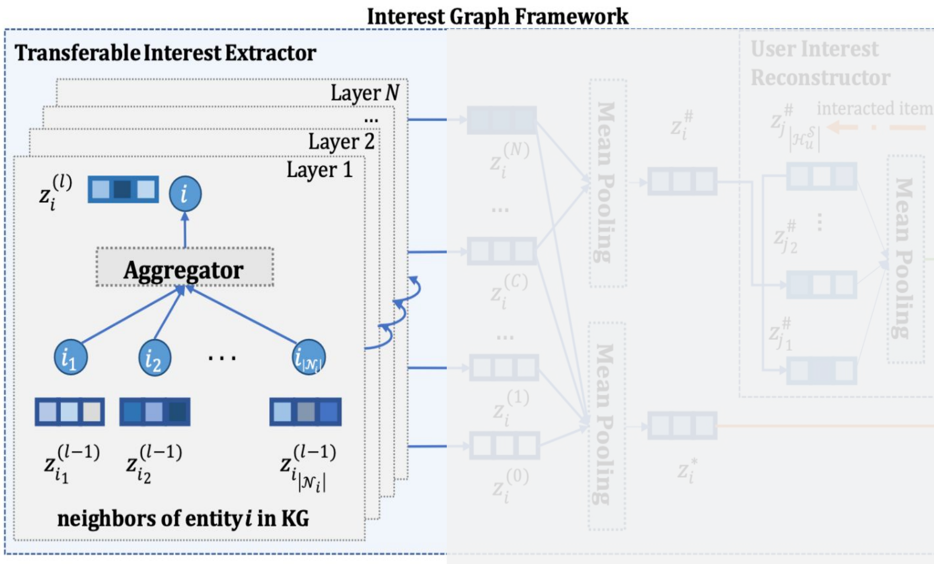
Method


Transferable interest extractor
User Interest Reconstructor
Domain Adaptation
Interest Graph Framework
Transferable Interest Extractor
Layer N
Layer 2
Layer 1
Aggregator
neighbors of entity i in KG
Mean Pooling
Mean Pooling
User Interest Reconstructor
interacted items
Mean Pooling
Trained in Source Domain
source domain
Tested in Target Domain
target domain

Transferable interest extractor

$$\mathbf{z}_i^{(0)} = \mathbf{e}_i \qquad \mathbf{e}_i \in \mathbb{R}^d$$

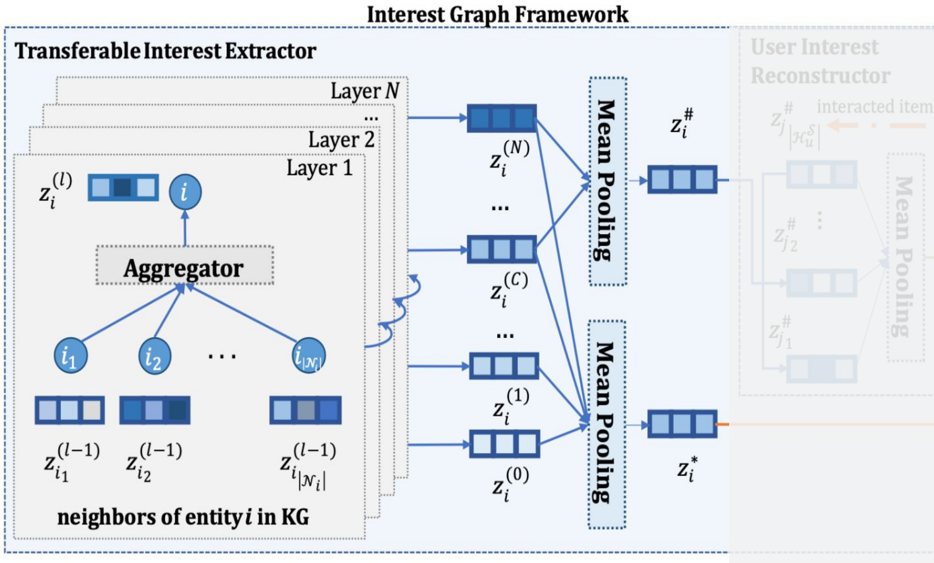
$$\mathbf{z}_i^{(l)} = \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \mathbf{z}_j^{(l-1)}$$



Transferable interest extractor

$$\mathbf{z}_i^* = \frac{1}{N+1}\sum_{j=0}^{N}\mathbf{z}_i^j$$

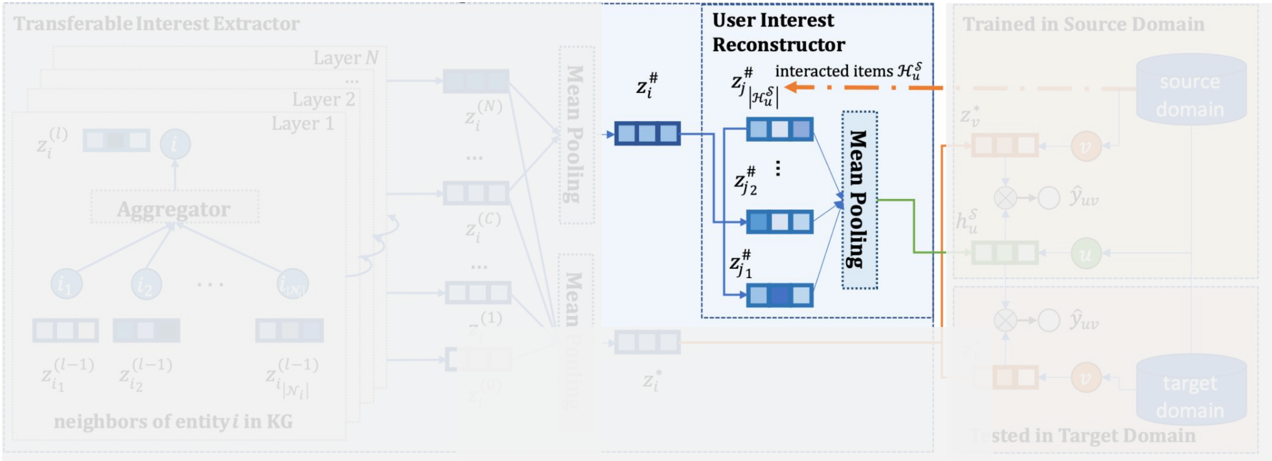
$$\mathbf{z}_i^\# = \frac{1}{N-C+1}\sum_{j=C}^{N}\mathbf{z}_i^j$$



User Interest Reconstructor

$$\mathbf{h}_u^S = \frac{1}{|\mathcal{H}_u^S|} \sum_{v \in \mathcal{H}_u^S} \mathbf{z}_v^\#$$

$$\mathbf{z}_i^\# = \frac{1}{N - C + 1} \sum_{j=C}^{N} \mathbf{z}_i^j$$

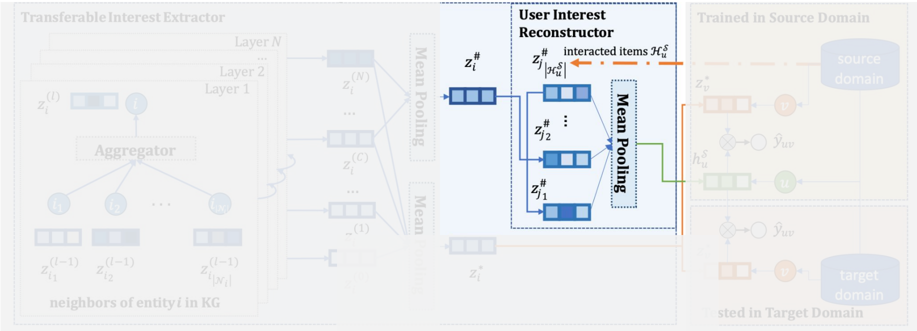


User Interest Reconstructor

$$\mathbf{z}_i^{\#} = \frac{1}{N - C + 1} \sum_{j=C}^{N} \mathbf{z}_i^{j}$$

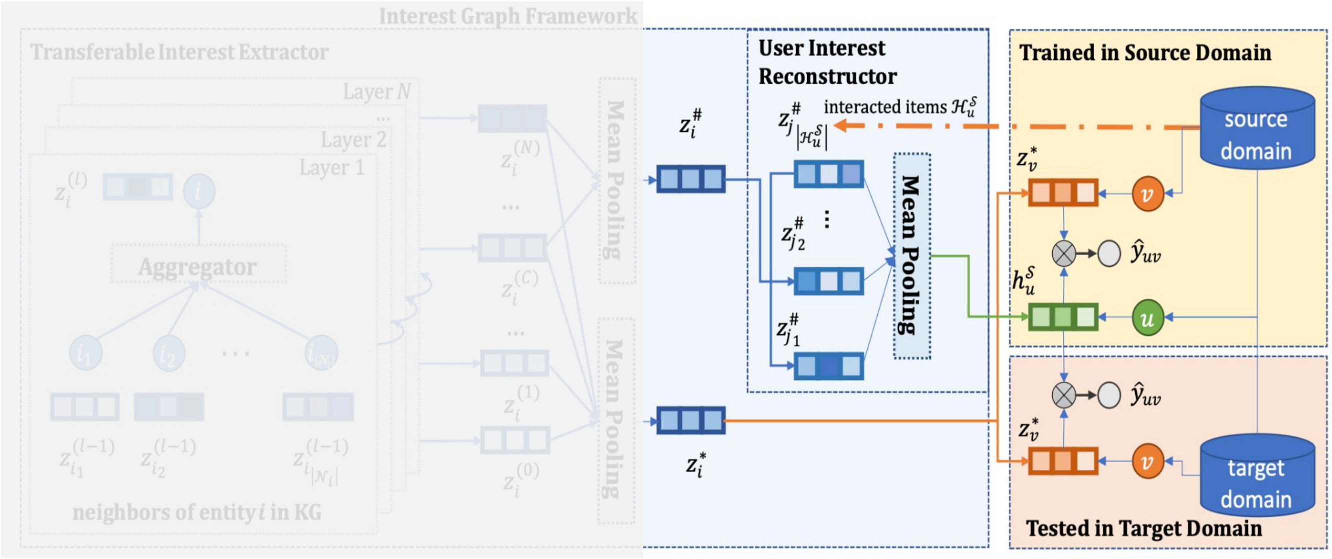
$$\mathbf{h}_u^{S} = \frac{1}{|\mathcal{H}_u^{S}|} \sum_{v \in \mathcal{H}_u^{S}} \mathbf{z}_v^{\#}$$

$$\mathbf{h}_u = \mathbf{e}_u$$



Domain Adaptation

$$\hat{y}_{uv} = \mathbf{z}_v^{*\top} \mathbf{h}_u^{\mathcal{S}}$$



Method

- **Loss function**: BPR (Bayesian Personalized Ranking from Implicit Feedback)

$$\mathcal{L} = -\sum_{u \in \mathcal{U}} \sum_{v \in \mathcal{H}_u^S} \sum_{v' \notin \mathcal{H}_u^S} \ln \sigma(\hat{y}_{uv} - \hat{y}_{uv'})$$

Outline

- Introduction
- Method
- **Experiment**
- Conclusion

Experiment

- **Dataset**

Dataset	AB	AM	ML	LFM
#User	11,240	11,240	6,040	18,029
#Item	47,377	16,100	3,655	311,994
#Interaction	202,223	142,395	997,580	1,006,639
#Entity	3,599,000			
#Relation	2,089			
#Triple	32,372,637			

Experiment

- **Daselines**
  - Random: lower bound
  - Oracle: upper bound（BPR）
  - NLP-based
  - KGE-based: transE
  - GCN-based: KGCN
  - Tiger(UserAsEmb): directly uses the ID-embedding to represent users

- **Evaluation**
  - Hit Ratio(H@K)
  - Normalization Discounted Cumulative Gain(N@K)

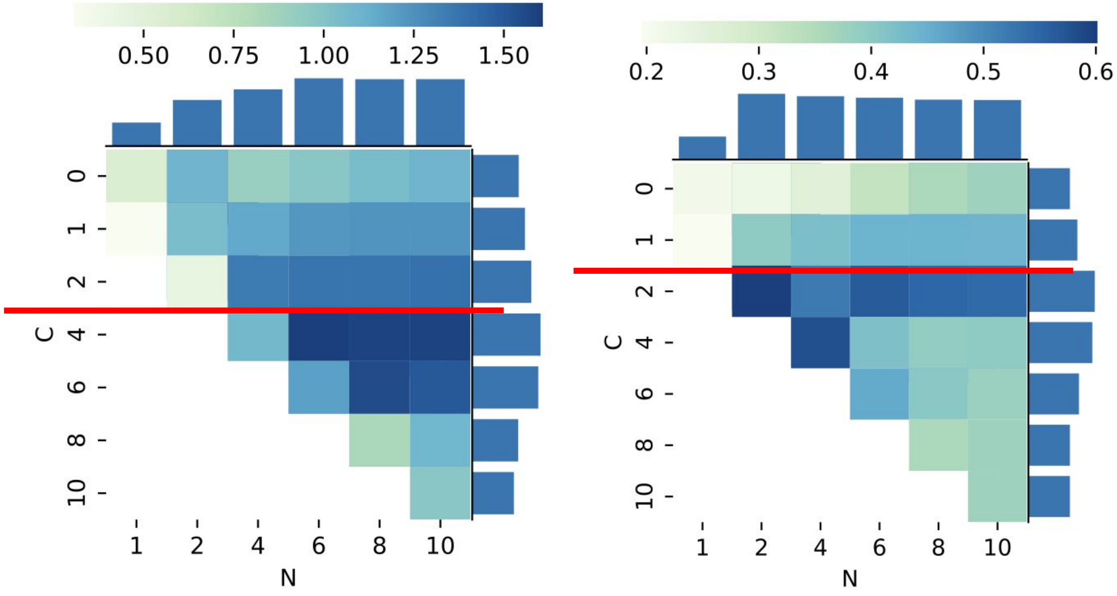
Experiment

Model	Source	Target	H@10	N@10	H@100	N@100
Random	-	AM	0.0620	0.0282	0.6211	0.1300
BPR (Oracle)	-	AM	2.8440*	1.4066*	14.1040*	3.5393*
NLP-based	AB	AM	0.1307	0.0488	1.2900	0.2586
TransE	AB	AM	0.3203	0.1580	1.4858	0.3719
KGCN	AB	AM	0.5368	0.2491	3.6032	0.8155
Tiger (UserAsEmb)	AB	AM	0.7711	0.4198	4.7242	1.1510
Tiger (normal)	AB	AM	0.9312	0.3751	**7.3072**	1.5401
Tiger (+ out domain)	ML+LFM+AB	AM	**1.0854**	**0.7484**	7.1886	**1.8811**
Random	-	AB	0.0211	0.0096	0.2111	0.0442
BPR (Oracle)	-	AB	0.7859*	0.4051*	3.9472*	1.0014*
NLP-based	AM	AB	0.0505	0.0202	0.4580	0.0948
TransE	AM	AB	0.0623	0.0293	0.3915	0.0913
KGCN	AM	AB	0.0860	0.0487	0.7117	0.1657
Tiger (UserAsEmb)	AM	AB	0.2343	0.1185	1.2604	0.3100
Tiger (normal)	AM	AB	0.3055	0.1370	1.9692	0.4519
Tiger (+ out domain)	ML+LFM+AM	AB	**0.5872**	**0.3392**	**2.5178**	**0.5659**

Experiment

N: number of GCN layer

C: number of discarded layer

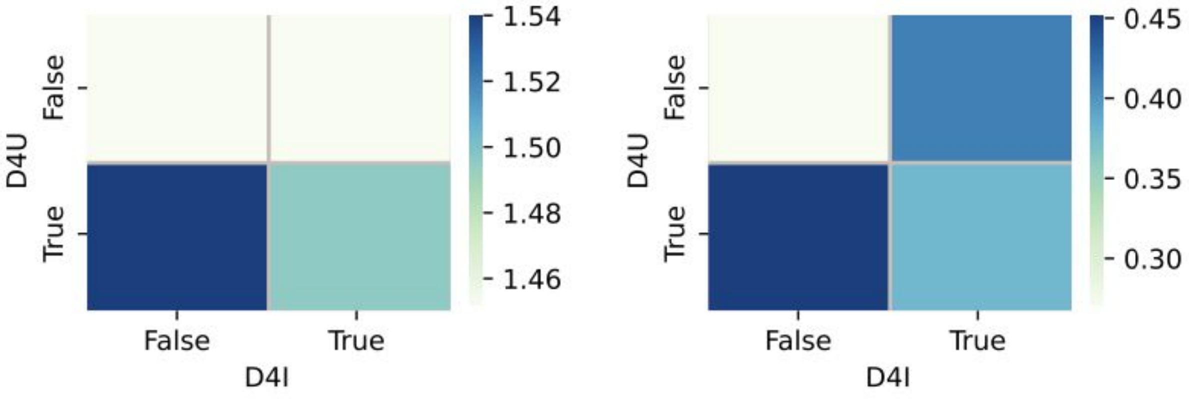


(a) AmazonMovie (b) AmazonBook

Experiment

D4U: use $z\#v$ or $z*v$ to reconstruct the user interest in Eq. (7)

D4I: use $z\#v$ or $z*v$ to present the item in Eq. (8)



(a) AmazonMovie (b) AmazonBook

Experiment

Training	Target	H@10	N@10	H@100	N@100
AB	AM	0.9312	0.3751	7.3072	1.5401
ML	AM	0.6139	0.2391	**7.3577**	1.4408
LFM	AM	0.9253	0.4784	6.3968	1.4474
ML+AB	AM	**1.1210**	0.5418	6.7438	1.6184
LFM+AB	AM	1.1121	0.3882	7.2598	1.4984
ML+LFM+AB	AM	1.0854	**0.7484**	7.1886	**1.8811**
AM	AB	0.3055	0.1370	1.9692	0.4519
ML	AB	0.3203	0.1456	1.7527	0.4121
LFM	AB	0.1275	0.0397	2.1501	0.4275
ML+AM	AB	0.5397	0.3013	2.0848	**0.6007**
LFM+AM	AB	0.4938	0.3211	1.6770	0.5434
ML+LFM+AM	AB	**0.5872**	**0.3392**	**2.5178**	0.5659

Conclusion

- They propose a solution “Tiger”, which project and fuse users’ universal preferences into a common interest graph bridging different domains’ collaborative behaviors.

- Develop better entity linking tools make more datasets can be linked to the interest graph, and perform interest graph pre-training to further shrink the gap between Tiger’s and the oracle model’s performance.

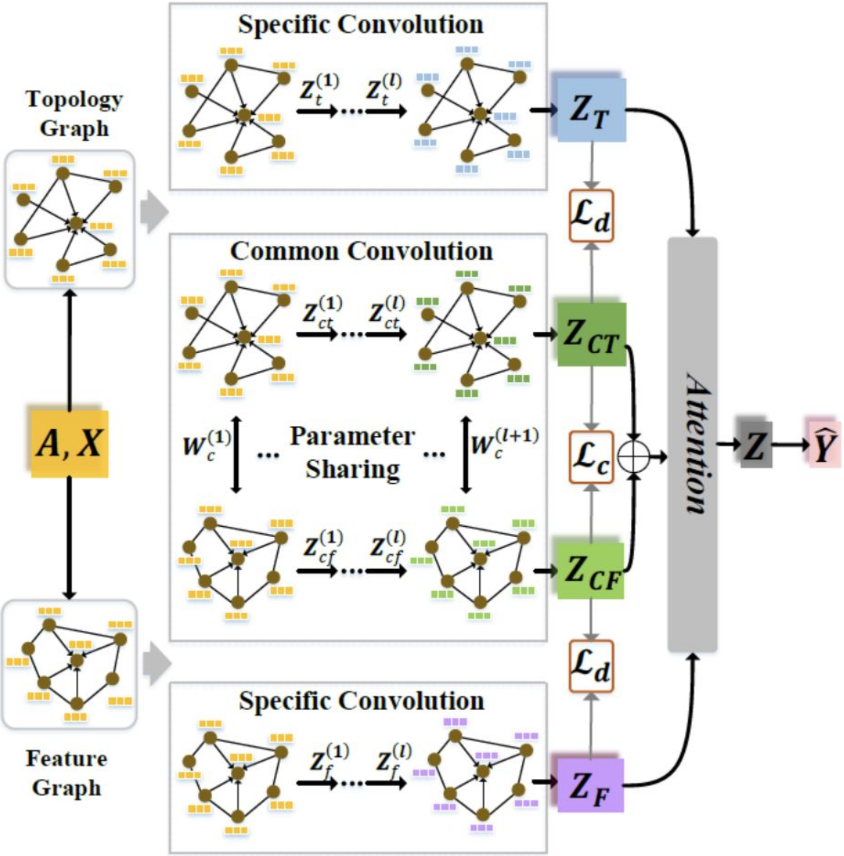
Introduction

- Domain-Level Zero-Shot
- Knowledge Graph
- GCN

Method

- Transferable Interest Extractor
- User Interest Reconstructor
- Domain Adaptation

Method


Specific Convolution
Topology Graph
$Z_t^{(1)}$ ... $Z_t^{(l)}$
Z_T
$\mathcal{L}_d$
Common Convolution
$Z_{ct}^{(1)}$ ... $Z_{ct}^{(l)}$
Z_{CT}
A, X
$W_c^{(1)}$ ... Parameter Sharing ... $W_c^{(l+1)}$
$\mathcal{L}_c$
Attention
Z
$\hat{Y}$
$Z_{cf}^{(1)}$ ... $Z_{cf}^{(l)}$
Z_{CF}
$\mathcal{L}_d$
Feature Graph
Specific Convolution
$Z_f^{(1)}$ ... $Z_f^{(l)}$
Z_F