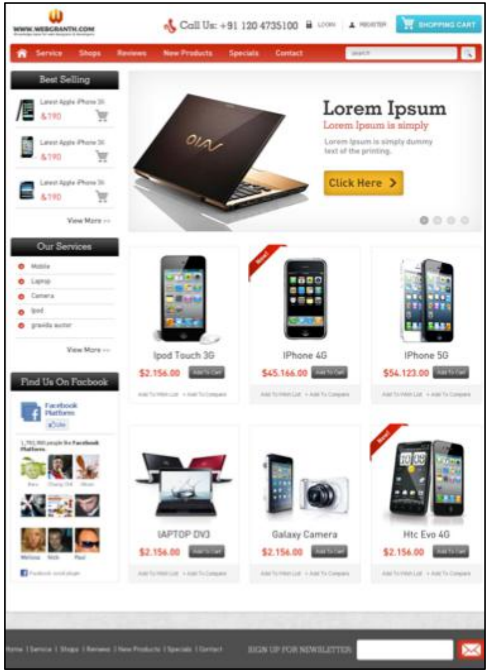
国际人工智能会议
AAAI 2021 论文北京预讲会

Graph Heterogeneous Multi-Relational Recommendation

Chong Chen, Min Zhang

Department of Computer Science and Technology,
Tsinghua University

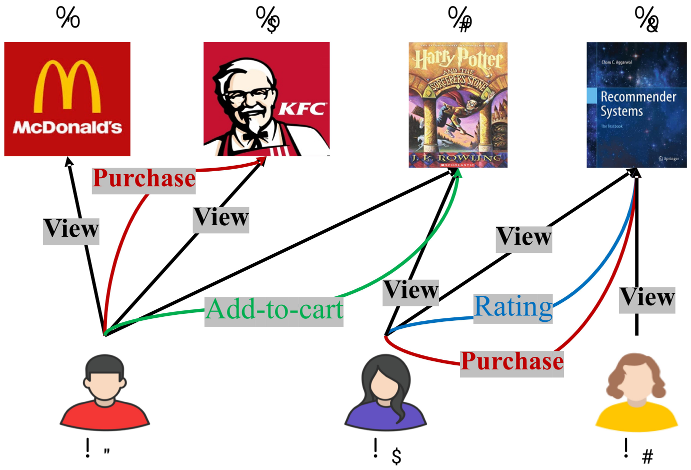
Background (1): Users' Sparse Feedback Information



- RecSys has become a major monetization tool for customer-oriented online services
  - E.g., E-commerce, News Portal, Social Networks, etc.
- Users usually rate or click a small set of items compared to hundreds of millions of items in the system
  - Sparse
  - Positive-unlabeled data

Background (2): Multi-relational Recommendation

- Multi-relational recommendation
  - Heterogeneous interactions between user and item (view, click, cart, purchase).
  - In addition to the target behavior, other behavior data also provides well-structured information and can be used for high-quality recommendation.



Preliminary:
Graph Convolutional Networks

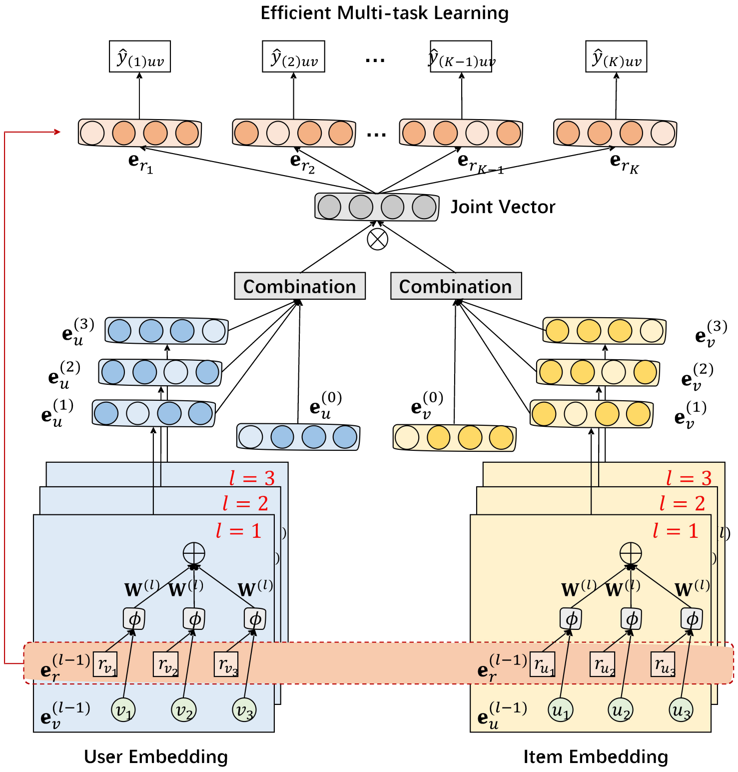
- Most existing research on graph convolutional networks are focused on learning representations of nodes in simple undirected graphs. $\mathbf{E} = \sigma(\hat{\mathbf{A}}\mathbf{E}^{(0)}\mathbf{W})$

- For capturing high-hop dependencies in the graph, several GCN layers can be stacke $\mathbf{E}^{(l)} = \sigma(\hat{\mathbf{A}}\mathbf{E}^{(l-1)}\mathbf{W}^{(l)})$

- For a relational graph $\mathbf{E}^{(l)} = \sigma(\hat{\mathbf{A}}\mathbf{E}^{(l-1)}\mathbf{W}_r^{(l)})$ ed GCN formulation is as follows

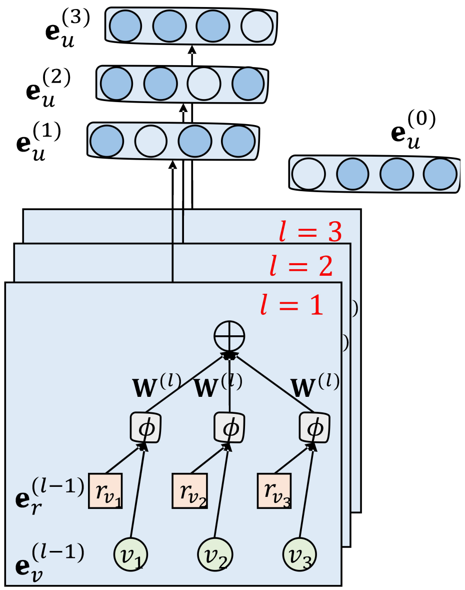
- This formulation leads to over-parameterization and embeds only nodes in the graph.

Graph Heterogeneous Collaborative Filtering



- A novel neural model named GHCF for multi-relational recommendation.
  - Relation-aware GCN propagation layers to jointly embed both representations of nodes and relations in a graph
  - Efficient non-sampling learning module to achieve more effective and stable model optimization
  - Multi-task prediction

Embedding Propagation Layers



- For node:

$$\mathbf{e}_u^{(l)} = \sigma\Big(\sum_{(v,r)\in\mathcal{N}(u)} \frac{1}{\sqrt{|\mathcal{N}_u|\,|\mathcal{N}_v|}} \mathbf{W}_r^{(l)} \mathbf{e}_v^{(l-1)} \Big)$$

$$\Downarrow$$

$$\mathbf{e}_u^{(l)} = \sigma\Big(\sum_{(v,r)\in\mathcal{N}(u)} \frac{1}{\sqrt{|\mathcal{N}_u|\,|\mathcal{N}_v|}} \mathbf{W}^{(l)} \phi(\mathbf{e}_v^{(l-1)}, \mathbf{e}_r^{(l-1)}) \Big)$$

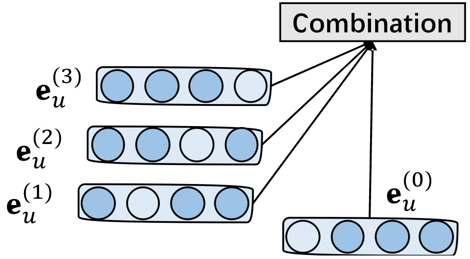
where

$$\phi(\mathbf{e}_v, \mathbf{e}_r) = \mathbf{e}_v \odot \mathbf{e}_r$$

- For relation:

$$\mathbf{e}_r^{(l)} = \mathbf{W}_{rel}^{(l)} \mathbf{e}_r^{(l-1)}$$

Multi-task Prediction

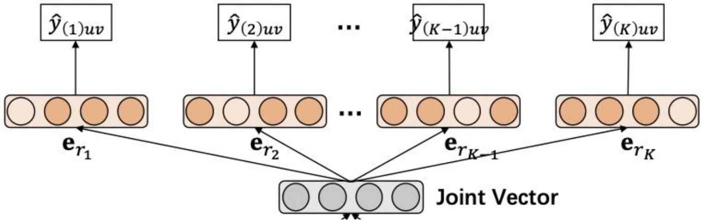


- Feature fusion

$$\mathbf{e}_u = \sum_{l=0}^{L} \frac{1}{L+1}\mathbf{e}_u^{(l)};\ \mathbf{e}_v = \sum_{l=0}^{L} \frac{1}{L+1}\mathbf{e}_v^{(l)};\ \mathbf{e}_r = \sum_{l=0}^{L} \frac{1}{L+1}\mathbf{e}_r^{(l)}$$

- Prediction

$$\hat{y}_{(k)uv} = \mathbf{e}_u^T \cdot diag(\mathbf{e}_{r_k}) \cdot \mathbf{e}_v = \sum_{i}^{d} e_{u,i} e_{r_k,i} e_{v,i}$$



Efficient Multi-task Learning without San

$$\tilde{\mathcal{L}}_k(\Theta) = \sum_{u\in\mathbf{B}} \sum_{v\in\mathbf{V}_{(u)}^{k+}} \left((c_v^{k+} - c_v^{k-})\hat{y}_{(k)uv}^2 - 2c_v^{k+}\hat{y}_{(k)uv} \right) + \sum_{i=1}^{d}\sum_{j=1}^{d}\left((e_{r_k,i} e_{r_k,j}) \left(\sum_{u\in\mathbf{B}} e_{u,i} e_{u,j}\right) \left(\sum_{v\in\mathbf{V}} c_v^{k-} e_{v,i} e_{v,j}\right)\right)$$

$$\mathcal{L}(\Theta) = \sum_{k=1}^{K} \lambda_k \tilde{\mathcal{L}}_k(\Theta) + \mu\|\Theta\|_2^2$$

Experimental settings

- Datasets:
- Baselines:
  - NCF (WWW 17)
  - ENMF (SIGIR 19)
  - LightGCN (SIGIR 20)
  - CMF (WWW 15)
  - MC-BPR (RecSys 16)
  - NMTR (ICDE 19)
  - EHCF (AAAI 19, our previous work)
- Evaluation methods: HR@K, NDCG@K, K=10, 50, 100

Dataset	#User	#Item	#View	#Add-to-cart	#Purchase
Beibei	21,716	7,977	2,412,586	642,622	304,576
Taobao	48,749	39,493	1,548,126	193,747	259,747

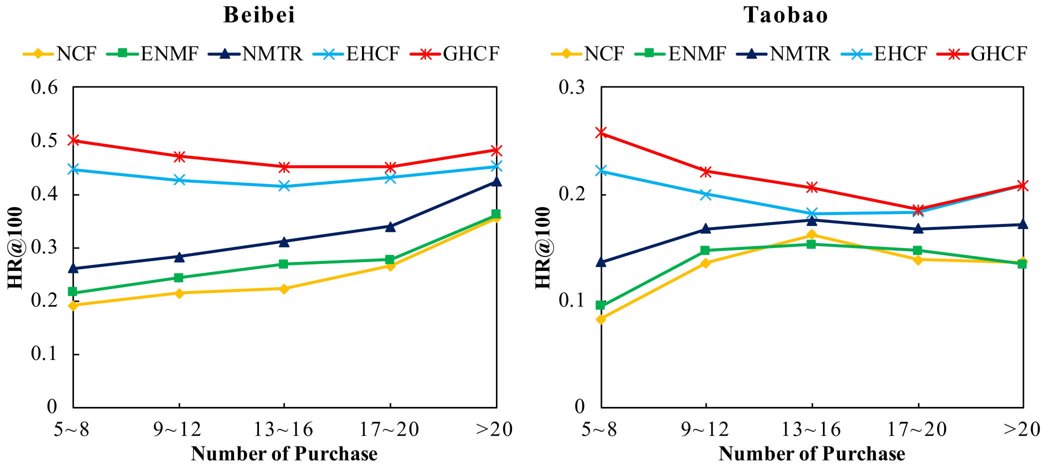
Model Comparisons

Beibei		HR@10	HR@50	HR@100	NDCG@10	NDCG@50	NDCG@100
Single	**BPR**	0.0437	0.1246	0.2192	0.0213	0.0407	0.0539
	NCF	0.0441	0.1562	0.2343	0.0225	0.0445	0.0584
	ENMF	0.0464	0.1637	0.2586	0.0247	0.0484	0.0639
	LightGCN	0.0451	0.1613	0.2495	0.0232	0.0466	0.0611
Heterogeneous	**CMF**	0.0482	0.1582	0.2843	0.0251	0.0462	0.0661
	MC-BPR	0.0504	0.1743	0.2755	0.0254	0.0503	0.0653
	NMTR	0.0524	0.2047	0.3189	0.0285	0.0609	0.0764
	EHCF	0.1523	0.3316	0.4312	0.0817	0.1213	0.1374
	GHCF	**0.1922****	**0.3794****	**0.4711****	**0.1012****	**0.1426****	**0.1575****

Taobao		HR@10	HR@50	HR@100	NDCG@10	NDCG@50	NDCG@100
Single	**BPR**	0.0376	0.0708	0.0871	0.0227	0.0269	0.0305
	NCF	0.0391	0.0728	0.0897	0.0233	0.0281	0.0321
	ENMF	0.0398	0.0743	0.0936	0.0244	0.0298	0.0339
	LightGCN	0.0415	0.0814	0.1025	0.0237	0.0325	0.0359
Heterogeneous	**CMF**	0.0483	0.0774	0.1185	0.0252	0.0293	0.0357
	MC-BPR	0.0547	0.0791	0.1264	0.0263	0.0297	0.0361
	NMTR	0.0585	0.0942	0.1368	0.0278	0.0334	0.0394
	EHCF	0.0717	0.1618	0.2211	0.0403	0.0594	0.0690
	GHCF	**0.0807****	**0.1892****	**0.2599****	**0.0442****	**0.0678****	**0.0792****

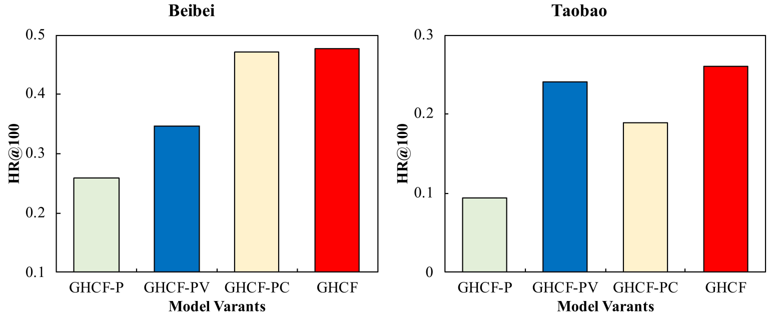
- Performance comparison on three datasets for all methods
- Best Baselines:
  - LightGCN: Single
  - EHCF: Heterogeneous
- GHCF
  - Consistently and significantly outperforms the best baseline

Handling Cold-Start Issue



- GHCF
  - Consistently significantly outperforms the best baseline
  - The effectiveness of leveraging auxiliary behavior to alleviate the data sparsity issue and the strong power of our GHCF model

Ablation Study



- GHCF-P: The variant model of GHCF which utilizes only purchase data.
- GHCF-PV: The variant model of GHCF which utilizes purchase data and view data.
- GHCF-PC: The variant model of GHCF which utilizes purchase data and carting data.

	Beibei		*Taobao*	
	HR@100	**NDCG@100**	**HR@100**	**NDCG@100**
GHCF-1	0.4569	0.1494	0.2473	0.0755
GHCF-2	0.4636	0.1498	0.2501	0.0778
GHCF-3	0.4674	0.1551	0.2567	0.0787
GHCF-4	0.4711	0.1575	0.2599	0.0792

Conclusion

Novel multi-task recommendation model

Relation-aware GCN propagation

Good performance

- We propose a novel neural model named GHCF for multi-relational recommendation, which uncovers the underlying relationships among heterogeneous user-item interactions and shows multi-task ability to predict various types of user behaviors using one unified model.
- We design relation-aware GCN propagation layers, which jointly embed both representations of nodes (users and items) and relations in a graph to explicitly exploit the collaborative high-hop signals.
- The model consistently outperforms the state-of-the-art recommendation methods, especially for cold-start users.

THANKS

2020.12.19