Local and Global Information Fusion for Top-N Recommendation in Heterogeneous Information Network

Binbin Hu[1], Chuan Shi[1], Wayne Xin Zhao[2], Tianchi Yang[1]

[1]Beijing University of Posts and Telecommunications

[2]Renming University of China

Background

Recommender System (RS)

- Discover items of interest from a large resource collection
- Basic recommendation method : Collaborative Filtering (e.g., matrix factorization)
- Suffer from cold-start problem
- Integrate more rich information (e.g., social network)

Heterogeneous Information Network (HIN)

- Include multiple types of nodes and links
- Model heterogeneous data and contain rich semantics
- Meta-path : Semantic paths between two objects in HIN

Local Information

- Direct interactions of users and items in HIN
- Breadth-first search (BFS)

Global Information

- Indirect interactions between users and items based on different meta-paths.
- Depth-first search (DFS)

Drawbacks of Traditional HIN based Recommendation

- Treat different local information equally
- seldom exploit and explore local information and global information simultaneously

Our idea

- Learn different weights for different neighbors
- An unified model to extensively exploit the local interaction information and fully explore the global interaction information

LGRec : The Proposed Method

Model Local Information

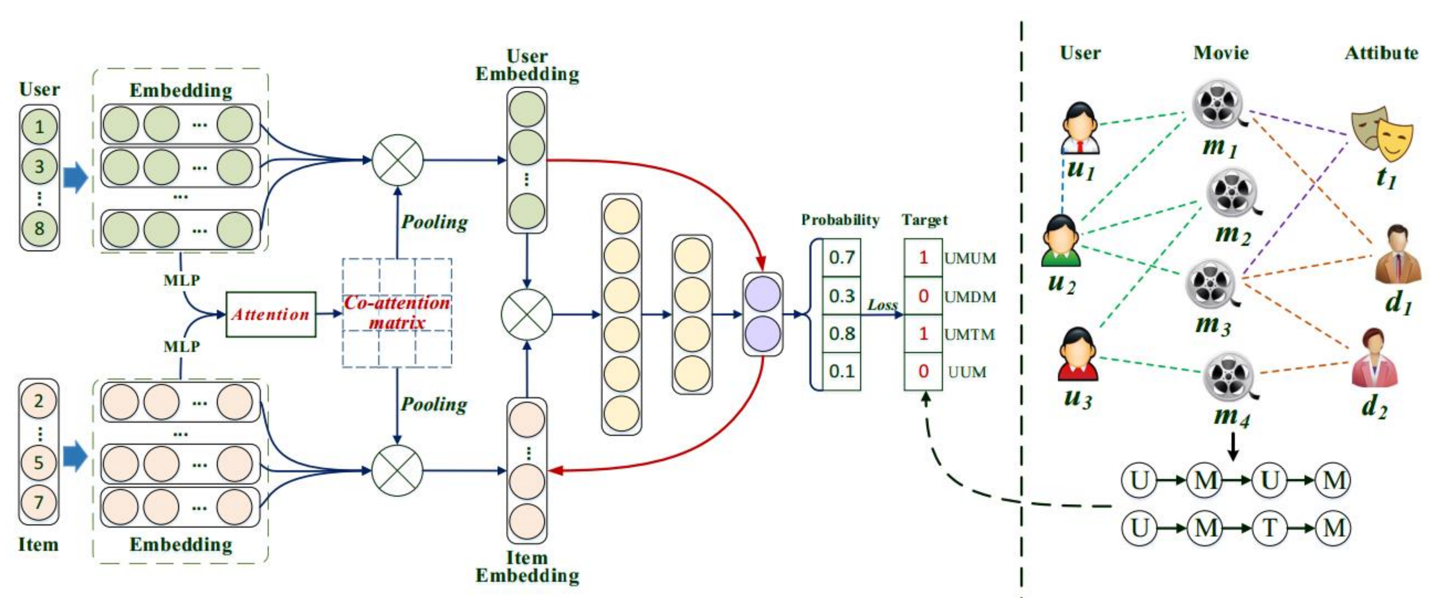
- Encoding user and item
- Co-attention mechanism

$$M_{i,j} = F(X_u^{(i)})AG(Y_v^{(j)})$$

- Generate embeddings

$$a_i^u = \mathrm{MP}(\{M_{ij}\}_{j=1}^{K_2}), \quad a_i^v = \mathrm{MP}(\{M_{ji}\}_{j=1}^{K_1}).$$

$$\boldsymbol{x}_u = \boldsymbol{X}_u \boldsymbol{a}^u, \quad \boldsymbol{y}_v = \boldsymbol{Y}_v \boldsymbol{a}^v.$$



Model Global Information

- Meta-path based interaction
- Generate latent relation based on MLP

$$\boldsymbol{h}_{u,v} = \boldsymbol{x}_u \oplus \boldsymbol{y}_v,$$
$$\boldsymbol{z} = \mathrm{MLP}(\boldsymbol{h}_{u,v}),$$

- Multi-label classification

$$\boldsymbol{p}_z = \boldsymbol{W}_o \boldsymbol{z} + \boldsymbol{b}_o,$$

$$\ell_{mc}(\boldsymbol{y}) = -\boldsymbol{y} * \log(\sigma(\boldsymbol{p}_z)) - (1-\boldsymbol{y}) * \log(1-\sigma(\boldsymbol{p}_z)) = \boldsymbol{p}_z - \boldsymbol{p}_z * \boldsymbol{y} + \log(1 + exp(-\boldsymbol{p}_z)).$$

Unified Model

$$s(u, v, z) = ||\boldsymbol{x}_u + \boldsymbol{z} - \boldsymbol{y}_v||_2^2.$$

$$\ell_{trans} = max(0, \lambda + s(u^+, v^+, z^+) - s(u^-, v^-, z^-)),$$

$$\ell = \ell_{trans} + \alpha[\ell_{mc}(\boldsymbol{y}^+) + \ell_{mc}(\boldsymbol{y}^-)] + \beta\ell_{reg}.$$

Performance

Datasets

Datasets	Relations (A-B)	#A	#B	#A-B	Meta-paths
Movielens	*User-Movie*	943	1,682	100,000	*UMUM*
	User-Age	943	8	943	*UMGM*
	User-Occupation	943	21	943	*UAUM*
	Movie-Genre	1,682	18	2,861	*UOUM*
LastFM	*User-Artist*	1,892	17,632	92,834	*UATA*
	User-User	1,892	1,892	18,802	*UAUA*
	Artist-Artist	17,632	17,632	153,399	*UUUA*
	Artist-Tag	17,632	11,945	184,941	*UUA*
Yelp	*User-Business*	16,239	14,284	198,397	*UBUB*
	User-User	16,239	16,239	158,590	*UBCaB*
	Business-City (Ci)	14,267	47	14,267	*UUB*
	Business-Category (Ca)	14,180	511	40,009	*UBCiB*
Amazon	*User-Item*	3,584	2,753	50,903	*UIUI*
	Item-View	2,753	3857	5,694	*UIVI*
	Item-Brand	2,753	334	2,753	*UIBI*
	Item-Category	2,753	22	5,508	*UICI*

Effectiveness Experiments

Models	Movielens		LastFM		Yelp		Amazon	
	HR@10	NDCG@10	HR@10	NDCG@10	HR@10	NDCG@10	HR@10	NDCG@10
ItemKNN	0.5854	0.3368	0.6327	0.5176	0.1480	0.0944	0.3153	0.1864
BRP	0.6766*	0.3860	0.7690	0.6061	0.5162	0.3186	0.3890	0.2195
MF	0.6702	0.3869*	0.7611	0.6051	0.5139	0.3176	0.3379	0.1893
NeuMF	0.6723	0.3816	0.7579	0.6070*	0.6660*	0.4218*	0.3619	0.2023
LRML	0.6140	0.3500	0.7204	0.5411	0.5934	0.3608	0.3304	0.1788
SVDFeature$_{hete}$	0.6033	0.3366	0.7848*	0.5813	0.6586	0.4117	0.3111	0.1575
FMG$_{rank}$	0.6267	0.3519	0.7758	0.5905	0.6080	0.3418	0.4154*	0.2244*
LGRec$_{noAtt}$	0.4836	0.2446	0.7104	0.5531	0.1372	0.0756	0.2720	0.1380
LGRec$_{noGlo}$	0.6564	0.3824	0.7717	0.6060	0.5894	0.3353	0.3820	0.2151
LGRec	**0.6914**	**0.3989**	**0.7865**	**0.6228**	**0.6902**	**0.4396**	**0.4235**	**0.2383**

Metrics

HR@10, NDCG@10

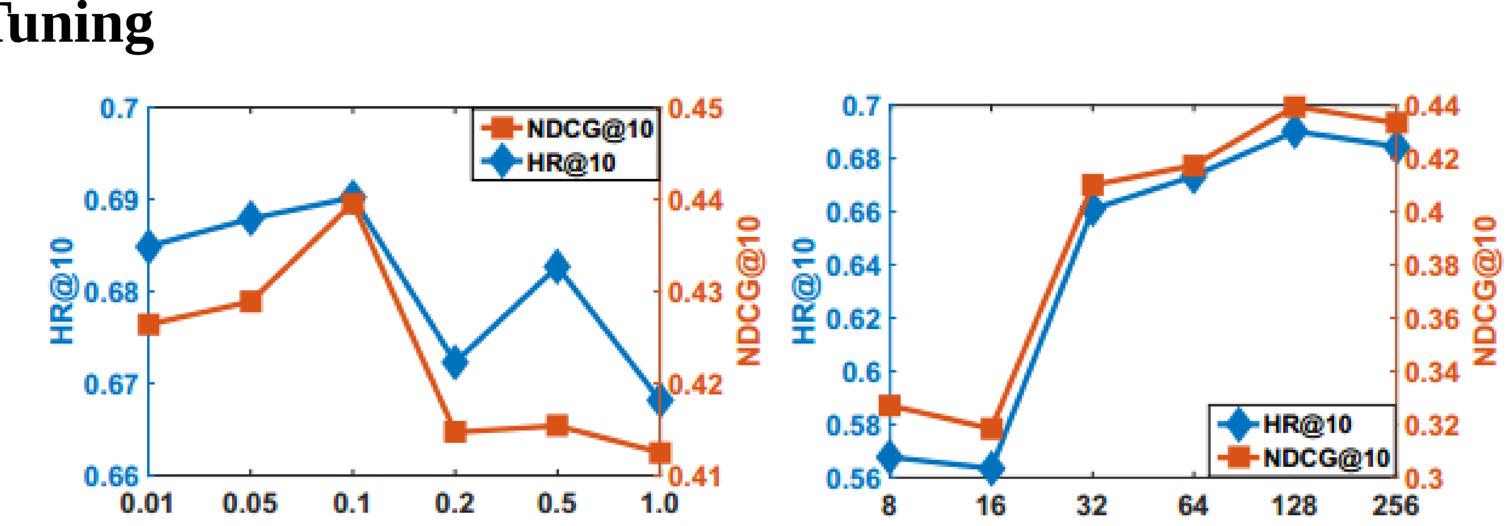
Compared Methods

CF based
- ItemKNN
- BPR
- MF
- NeuMF
- LRML

HIN based
- SVDFeature$_{hete}$
- FMG$_{rank}$

Ours
- LGRec$_{noAtt}$
- LGRec$_{noGlo}$
- LGRec

Parameters Tuning



(a) Varying α (b) Varying d

Conclusions

- We proposed a unified deep model to fully utilize local and global information for top-N recommendation.
- We learns importance of neighbors by the co-attention mechanism and optimize a multi-label classification problem to capture meta-path based interactions.
- Extensive experimental results show the effectiveness of LGRec.
- More materials in webpage : www.shichuan.org

Acknowledgements

This work is supported in part by the National Natural Science Foundation of China (No. 61772082, 61502502, 61320106006, 61375058), the National Key Research and Development Program of China (2017YFB0803304), and the Beijing Municipal Natural Science Foundation (4182043, 4162032).

Contact

- hubinbin@bupt.edu.cn
- shichuan@bupt.edu.cn
- batmanfly@gmail.com
- yangtianchi@bupt.edu.cn