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

Binbin Hu[†], Chuan Shi[†*], Wayne Xin Zhao[‡], Tianchi Yang[†]

[†] Beijing University of Posts and Telecommunications, Beijing, China
[‡] School of Information, Renmin University of China, Beijing, China
{hubinbin,shichuan,yangtianchi}@bupt.edu.cn,batmanfly@gmail.com

ABSTRACT

Since heterogeneous information network (HIN) is able to integrate complex information and contain rich semantics, there is a surge of HIN based recommendation in recent years. Although existing methods have achieved performance improvement to some extent, they still face the following problems: how to extensively exploit and comprehensively explore the local and global information in HIN for recommendation. To address these issues, we propose a unified model LGRec to fuse local and global information for top-N recommendation in HIN. We firstly model most informative local neighbor information for users and items respectively with a co-attention mechanism. In addition, our model learns effective relation representations between users and items to capture rich information in HIN by optimizing a multi-label classification problem. Finally, we combine the two parts into an unified model for top-N recommendation. Extensive experiments on four real-world datasets demonstrate the effectiveness of the proposed model.

KEYWORDS

Heterogeneous Information Network, Recommender System, Local and Global Information, Attention Mechanism

ACM Reference Format:
Binbin Hu, Chuan Shi, Wayne Xin Zhao, Tianchi Yang. 2018. Local and Global Information Fusion for Top-N Recommendation in Heterogeneous Information Network. In The 27th ACM International Conference on Information and Knowledge Management (CIKM'18), October 22–26, 2018, Torino, Italy. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3269206.3269278

1 INTRODUCTION

In the era of information explosion, recommender systems have been playing a pivotal role in various online services [5]. Classic recommendation methods, e.g., matrix factorization [4], mainly model users' preference towards items using historical user-item interaction records. Nowadays, various kinds of auxiliary data become available in recommender systems, which can be leveraged to improve recommendation performance.

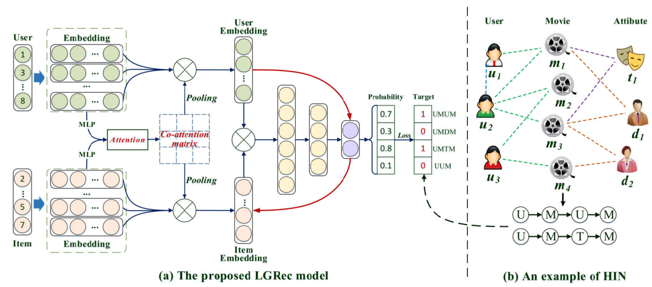


Figure 1: The overall architecture of the proposed model.

Recently, heterogeneous information network (HIN) , consisting of either multiple types of nodes or links, has been proposed as a powerful modeling method to fuse complex information, and is successfully applied to many data mining tasks [7]. Moreover, meta-path [7], a relation sequence connecting objects, is widely used to effectively explore rich information and network structure in HIN. Due to its flexibility in modeling data heterogeneity, HIN has also been adopted in recommender systems to characterize rich auxiliary data in recent years, and those algorithms are also called HIN based recommendation methods. In Fig. 1(b), we present an example for movie recommendation characterized by HIN. We can see that the HIN contains multiple types of entities connected by different types of relations. And the meta-path $User - Movie - User - Movie$ (UMUM) indicates the historical interaction records between users and movies, while $User - Movie - Type - Movie$ (UMTM) indicates users prefer movies with the same type. Based on HIN, we further explore useful information for recommendation, namely *local information* and *global information*. Concretely, local information is direct interactions of users and items in HIN, while global information is indirect interactions between users and items based on different meta-paths. As we can see in Fig. 1(b), u_3 directly interacts with m_2 and m_4, which can be considered as the local information of u_3. Besides, u_3 can interact with m_3 through path $u_3 - m_2 - u_2 - m_3$ (UMUM) or path $u_3 - m_4 - d_2 - m_3$ (UMDM), which are the global information.

Existing HIN based recommendation methods [6, 9] usually leverage path based semantic relatedness of user-item pairs to enhance the representations of users and items for recommendation.

*Corresponding Author.

CIKM '18, October 22–26, 2018, Torino, Italy

ACM ISBN 978-1-4503-6014-2/18/10…$15.00
https://doi.org/10.1145/3269206.3269278

Although these HIN based methods have achieved performance improvement to some extent, there are two shortcomings: (1) These methods tend to treat different local information equally, which is not an effective way to characterize these information for recommendation. (2) They seldom exploit and explore local information and global information simultaneously. In HIN, besides the direct interactions (local information) between users and items, there widely exist meta-path based interactions (global information), which can potentially be integrated for recommendation. Based on these considerations, we aim to propose an unified model to extensively exploit the local interaction information and fully explore the global interaction information for top-N recommendation.

In order to comprehensively utilize local information, we assume that the embedding of a user (item) is determined by its connected items (users). Since different items (users) may have different contributions to the users (items), a co-attention mechanism is designed to automatically determine their weights. And then we explore and exploit rich information and network structure in HIN which is the global information can be modeled as the relation of each user-item pair. The learned representations can be regard as the composite interactions between users and items based on multiple meta-paths. Therefore a multi-label classification task is modeled to match the relation embdedding and interaction distribution on meta-paths. Furthermore, we design a joint optimization objective to model the translation mechanism among users', items' representation and corresponding relation representation as well as the multi-label classification task for top-N recommendation. Extensive experiments on four real-world datasets demonstrate the effectiveness of the proposed model compared to the state of arts.

2 THE PROPOSED MODEL

In this paper, we present an unified model to fuse *L*ocal and *G*lobal information for top-*N* *Rec*ommendation in heterogenous information network, called **LGRec**. We present the overall architecture for the proposed model in Fig. 1(a). As we can see, our model flexibly utilizes the local neighbor information and apply a co-attention mechanism to construct more meaningful embeddings for users and items. And then we learn the relation representation through interactions between users and items based on MLP. Moreover, we utilize the learned relation to predict the meta-path based interaction distribution in HIN, which is modeled as a multi-label classification problem. Finally, we combine the two parts in an unified model and optimize it for top-N recommendation.

2.1 Local Information based Recommendation Model with Co-attention Mechanism

First of all, we propose a basic local information based recommendation model, which learn embeddings of users (items) according to that of connecting items (users) with a co-attention mechanism.

Encoding user and item. Each user (item) can be represented as a sequence of items (users) which have been interacted, *i.e.,* $\boldsymbol{p}_u \in \mathbb{R}^{K_1 \times 1}$ and $\boldsymbol{q}_v \in \mathbb{R}^{K_2 \times 1}$, where K_1 and K_2 are the number of neighbors of user u and item v representatively. Due to the unbalanced degree distribution of a vertex, we propose to utilize a MF based model [4] for ranking direct neighbors, and keep the top K_1 and K_2 neighbors as users' and items' neighborhood respectively. Following [2], we set up lookup layers to transform each user and item into low-dimensional dense vector. Therefore, we encode local information of user u and item v into $\boldsymbol{X}_u \in \mathbb{R}^{K_1 \times d}$ and $\boldsymbol{Y}_v \in \mathbb{R}^{K_2 \times d}$, where d is the dimension of embedding.

Co-attention mechanism. Since different items (users) have different importance to a user (item), we cannot treat them equally. And thus we propose to select the most informative local information for each user and item respectively and generate more meaningful representations of users and items. Given the local information embedding matrices of a user $\boldsymbol{X}_u \in \mathbb{R}^{d \times K_1}$ and an item $\boldsymbol{Y}_v \in \mathbb{R}^{d \times K_2}$, we calculate a co-attention matrix $\boldsymbol{M} \in \mathbb{R}^{K_1 \times K_2}$ between them. Each entry of $\boldsymbol{M}$ can be described as follows:

$$\boldsymbol{M}_{i,j} = F(X_u^{(i)}) A G(Y_v^{(j)}), \tag{1}$$

where $\boldsymbol{A}$ is an attentive matrix. $F(\cdot)$ and $G(\cdot)$ are neural network functions with multiple layers for users and items respectively.

Generate embeddings. After that, we conduct max pooling (MP) operations along rows and columns of $\boldsymbol{M}$ to generate the importance vectors for users and items respectively, which can be describe as follows:

$$\boldsymbol{a}_i^u = \mathrm{MP}(\{\boldsymbol{M}_{ij}\}_{j=1}^{K_2}), \quad \boldsymbol{a}_i^v = \mathrm{MP}(\{\boldsymbol{M}_{ji}\}_{j=1}^{K_1}). \tag{2}$$

Next, we employ softmax function to normalize the above importance vectors and aggregate neighborhood information for final embeddings as follows:

$$\boldsymbol{x}_u = \boldsymbol{X}_u \boldsymbol{a}^u, \quad \boldsymbol{y}_v = Y_v \boldsymbol{a}^v. \tag{3}$$

Basic recommendation model. Motivated by the translation mechanism in collaborative filtering [8], we preserve the translation mechanism among user and item representations. Hence, for each user-item pair $\langle u, v \rangle$, we define the scoring function as follows:

$$s(u, v) = ||\boldsymbol{x}_u - \boldsymbol{y}_v||_2^2. \tag{4}$$

2.2 Modeling Global information with Multi-label Classification

Since each user and item pair can be connected by multiple meta-paths, the relation of a user and an item is the combination of composite interaction based on these meta-paths. We believe that the learned relation representation can not only capture the rich information and network structure in HIN, but also further predict the meta-path based interaction distribution between users and items. And thus, we can model the problem as the multi-label classification and integrate it into final unified objective.

Meta-path based interaction. In order to model the global information, we firstly obtain the interactions between the source and the target nodes based on meta-paths. Supposed we have a meta-path $\rho = (A_1, A_2, \ldots, A_l)$, where A_i represents the node type. Then we can define a matrix $\boldsymbol{C}_{A_i A_j}$ as the adjacency matrix between type A_i and type A_j. Then, we define the interaction matrix for meta-path ρ is $\boldsymbol{I}^\rho = \boldsymbol{C}_{A_1 A_2} \circ \boldsymbol{C}_{A_2 A_3} \circ \ldots \circ \boldsymbol{C}_{A_{l-1} A_l}$. Each entry of matrix $\boldsymbol{I}_{ij}^\rho$ represents whether there exist any interaction between the source node i and the target node j based on meta-path ρ.

Generating latent relation based on MLP. Given the user-item pair $\langle u, v\rangle$, our model firstly applies the following step to learn a joint embedding of users and items:

$$\boldsymbol{h}_{u,v} = \boldsymbol{x}_u \oplus \boldsymbol{y}_v, \tag{5}$$

where "$\oplus$" denotes the vector concatenation operation. We aim to generate latent relation for a user and an item through the composite interaction between them. Following [2], we feed $\boldsymbol{h}_{u,v}$ into a MLP component in order to implement a nonlinear function for modeling complicated the latent relation.

$$\boldsymbol{z} = \mathrm{MLP}(\boldsymbol{h}_{u,v}), \tag{6}$$

where the MLP component is implemented with two hidden layers with ReLU as the activation function.

Multi-label classification. Since we obtain the relation representation (*i.e.,* $\boldsymbol{z}$) for a user-item pair (*i.e.,* $\langle u, v\rangle$), we could leverage the output vector to predict the probability of interactions between u and v based on each meta-path $\rho \in \mathcal{P}$. The probability can be generated as follows:

$$\boldsymbol{p_z} = \boldsymbol{W_o z} + \boldsymbol{b_o}, \tag{7}$$

where $\boldsymbol{W_o} \in \mathbb{R}^{|\mathcal{P}|\times d}$ and $\boldsymbol{b_o} \in \mathbb{R}^{|\mathcal{P}|\times 1}$ are the weight matrix and bias respectively. $\boldsymbol{p_z}$ is the prediction vector of length $|\mathcal{P}|$ consisting of the probability of each meta-path based interaction between u and v. Then we encode the interaction distribution based on multiple meta-paths between users and item as one-hot vector, denoted as $\boldsymbol{y} \in \mathbb{R}^{|\mathcal{P}|\times 1}$. Since we obtain the interaction matrix $\boldsymbol{I}^{\rho}$ for each meta-path ρ as mentioned above, we formally define $\boldsymbol{y} = \{I_{uv}^{\rho} | \rho \in \mathcal{P}\}$ when giving a user u and an item v. In other words, each entry of $\boldsymbol{y}$ represents whether corresponding meta-path based interaction exists between user u and item v. We believe that the learned relation representation can predict the meta-path based interaction distribution between users and items, which can be naturally modeled as a multi-label classification problem. Hence, we use the sigmoid cross entropy with logits as our objective to overcome this issue, which can be described as follows:

$$\begin{aligned} \ell_{mc}(\boldsymbol{y}) &= -\boldsymbol{y} * \log(\sigma(\boldsymbol{p_z})) - (1-\boldsymbol{y}) * \log(1-\sigma(\boldsymbol{p_z})) \\ &= \boldsymbol{p_z} - \boldsymbol{p_z} * \boldsymbol{y} + \log(1 + exp(-\boldsymbol{p_z})). \end{aligned} \tag{8}$$

2.3 Unified Model

Following [8], we consider the global heterogeneous information based relation into recommendation. For each user-item pair $\langle u, v\rangle$, we learn the HIN based relation $\boldsymbol{z}$, and we extend the original scoring function (Eq. 4) as:

$$s(u, v, \boldsymbol{z}) = ||\boldsymbol{x}_u + \boldsymbol{z} - \boldsymbol{y}_v||_2^2. \tag{9}$$

And then, we adopt the hinge loss for optimization. For each positive user-item pair $\langle u^+, v^+\rangle$, we sample a negative pair denoted as $\langle u^-, v^-\rangle$. As mentioned above, we can learn corresponding relation representation $\boldsymbol{z}^+$ and $\boldsymbol{z}^-$, respectively. The hinge loss is defined as follows:

$$\ell_{trans} = max(0, \lambda + s(u^+, v^+, \boldsymbol{z}^+) - s(u^-, v^-, \boldsymbol{z}^-)), \tag{10}$$

where $\lambda > 0$ is the margin hyper-parameter.

To preserve the translation mechanism among user's and item's representation, and match relation representation with interaction distribution based on meta-paths, we combine the objective in Eq. 8 and Eq. 10 and proposed an unified recommendation model. For each $\langle u^+, v^+, \boldsymbol{y}^+\rangle$ and its negative sample $\langle u^-, v^-, \boldsymbol{y}^-\rangle$, the model jointly optimizes the objective as follows:

$$\ell = \ell_{trans} + \alpha[\ell_{mc}(\boldsymbol{y}^+) + \ell_{mc}(\boldsymbol{y}^-)] + \beta\ell_{reg}. \tag{11}$$

Here, we introduce two hyper-parameters α and β to balance the weights of different parts. Besides, ℓ_{reg} is an L_2-norm regularizer to prevent overfitting. To optimize our model, we perform minibatch Adam with a batch size of 256. The learning rate is tuned amongst {0.0001, 0.0005, 0.001}. The margin λ is tuned amongst {0.1, 0.5, 1.0, 2.0}. And we set $K_1 = K_2 = 100$, $\alpha = 0.1$, $\beta = 0.001$ and $d = 128$. In addition, we use the meta-paths reported in Table 1.

Table 1: Statistics of the four datasets. The last column reports the selected meta-paths in each dataset.

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*

3 EXPERIMENTS

Datasets and evaluation metrics. We evaluate our proposed model over four real-world datasets, namely MovieLens , LastFM , Yelp [3] and Amazon[1], The detailed descriptions of the four datasets are summarized in Table 1. We adopt the widely used leave-one-out method to evaluate the performance of item recommendation [2, 8]. For each golden item (*e.g.,* the item user has interacted with) in the test, we randomly sample 100 negative items and rank the golden item among the 100 items. Following [2], we adopt Hit ratio at rank k (HR@k) and Normalized Discounted Cumulative Gain at rank k (NDCG@k) to evaluate the performance of model performance.

Compared methods. We consider the following representative recommendation methods for performance comparisons. (1) *ItemKNN*: It is a typical collaborative filtering method by recommending similar items based on past items [5]; (2) *BPR*: It optimizes the MF model with the pairwise ranking loss [4]; (3) *MF*: It optimizes the classical MF model with the cross entropy loss [2]; (4) *NeuMF*: It is the recently proposed neural network method for top-N recommendation [2]; (5) *LRML*: It is the state-of-the-art memory based attention model for item recommendation [8]; (6) $SVDFeature_{hete}$: We extract heterogeneous information as one-hot feature to feed into *SVDFeature* for recommendation [1]; (7) FMG_{rank}: It is the state-of-the-art HIN based recommendation by optimizing the pair ranking loss [9].

[1]http://jmcauley.ucsd.edu/data/amazon/

Table 2: **Results of effectiveness experiments on four datasets. We use "*" to mark the best performance from the baselines.**

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**

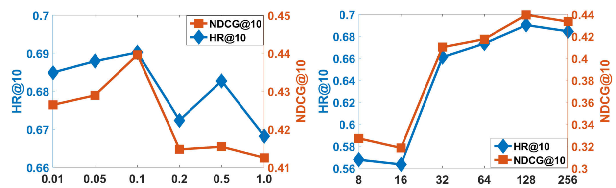
To examine the effectiveness of the co-attention mechanism and the global information modeling method, we consider two variants as compared baselines, namely LGRec$_{noAtt}$ (our model without co-attention mechanism) and LGRec$_{noGlo}$ (our model without global information). And LGRec is our complete model.

Results and analysis. We report the comparison results of our proposed model and baselines on four datasets in Table 2. The major findings from the experimental results are summarized as follows: (1) LGRec is consistently better than all the baselines on the four datasets. This observation demonstrates the effectiveness of our model on the task of top-N recommendation, which is more capable of utilizing local neighborhood information and global heterogenous information. (2) Considering the two variants of LGRec, we can find that the overall performance order is as follows: LGRec > LGRec$_{noGlo}$ > LGRec$_{noAtt}$. The result indicates that the gobal information and co-attention mechanism really work in our model and the global information play a critical role for the performance improvement in most cases. (3) Among these baselines, HIN based methods (SVDFeature$_{hete}$ and FMG$_{rank}$) outperform CF methods (ItemKNN, BPR and MF) in most cases, which indicates the usefulness of heterogeneous information. In addition, NeuMF also achieves competitive performance due to the adoption of neural network, while its performance is still worse than LGRec because of the absence of heterogeneous information.

Parameter tuning. We examine the effect of balance parameter α and the dimension of embeddings d on the performance for our model. As shown in Fig 2, we can see that (1) when $\alpha = 0.1$, our model achieves the best performance, indicating that the balance parameter should be set to a small number; and (2) our model achieves the best performance when $d = 128$, which indicates that the dimension of embeddings cannot be set too small or too large.

4 CONCLUSION

In this paper, we proposed a novel deep neural network model to fully utilize local and global information for top-N recommendation in HIN. The model firstly learns user (item) embeddings according to the neighbor items (users) with a co-attention mechanism. In addition, our model learns relation representations between users and items to capture meta-path based interactions by optimizing a multi-label classification problem. Considering these two factors, an unified optimization objective is learned for top-N recommendation. Extensive experimental results have demonstrated the recommendation effectiveness of our model.



(a) Varying α (b) Varying d

Figure 2: Performance tuning with the varying of the parameter α and the dimension of embeddings d on Yelp dataset.

5 ACKNOWLEDGEMENT

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).

REFERENCES

[1] Tianqi Chen, Weinan Zhang, Qiuxia Lu, Kailong Chen, Zhao Zheng, and Yong Yu. 2012. SVDFeature: a toolkit for feature-based collaborative filtering. *JMLR* 13, Dec (2012), 3619–3622.

[2] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In *WWW*. 173–182.

[3] Binbin Hu, Chuan Shi, Wayne Xin Zhao, and Philip S Yu. 2018. Leveraging Meta-path based Context for Top-N Recommendation with A Neural Co-attention Model. In *SIGKDD*. 1531–1540.

[4] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *UAI*. 452–461.

[5] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *WWW*. 285–295.

[6] Chuan Shi, Binbin Hu, Wayne Xin Zhao, and Philip S Yu. 2018. Heterogeneous Information Network Embedding for Recommendation. *TKDE* (2018).

[7] Chuan Shi, Yitong Li, Jiawei Zhang, Yizhou Sun, and S Yu Philip. 2017. A survey of heterogeneous information network analysis. *TKDE* 29, 1 (2017), 17–37.

[8] Yi Tay, Luu Anh Tuan, and Siu Cheung Hui. 2018. Latent Relational Metric Learning via Memory-based A ention for Collaborative Ranking. In *WWW*. 729–739.

[9] Huan Zhao, Quanming Yao, Jianda Li, Yangqiu Song, and Dik Lun Lee. 2017. Meta-graph based recommendation fusion over heterogeneous information networks. In *SIGKDD*. 635–644.