Leveraging Meta-path based Context for Top-N Recommendation with a Neural Co-attention Model

Binbin Hu[1], Chuan Shi[1], Wayne Xin Zhao[2], Philip S. Yu[3]
[1]Beijing University of Posts and Telecommunications, Beijing, China
[2]Renming University of China, Beijing, China
[3]University of Illinois at Chicago, IL, USA

KDD2018

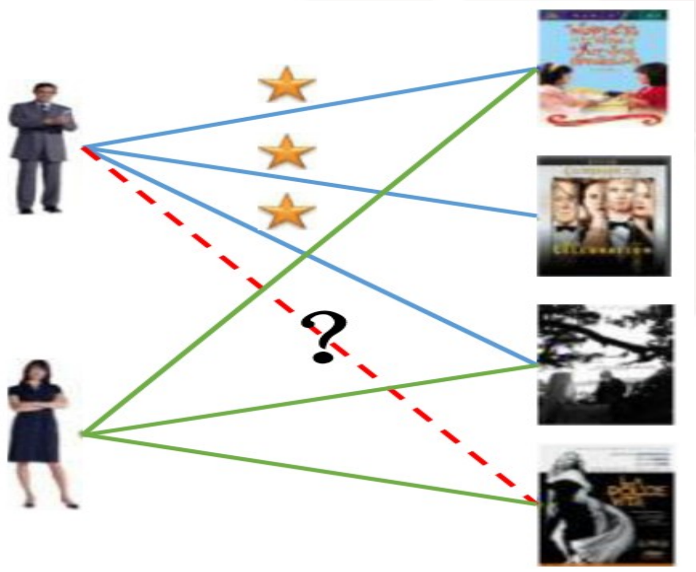
CONTENTS

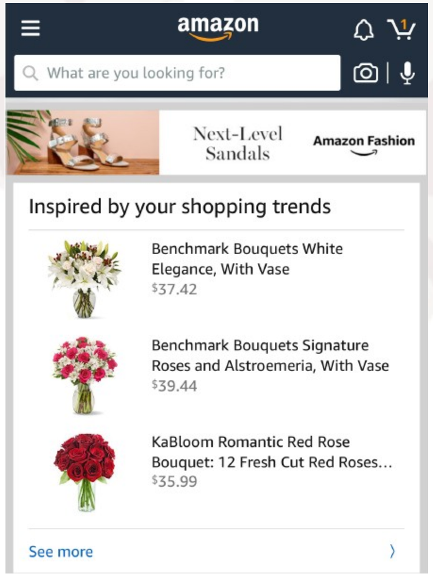
KDD2018

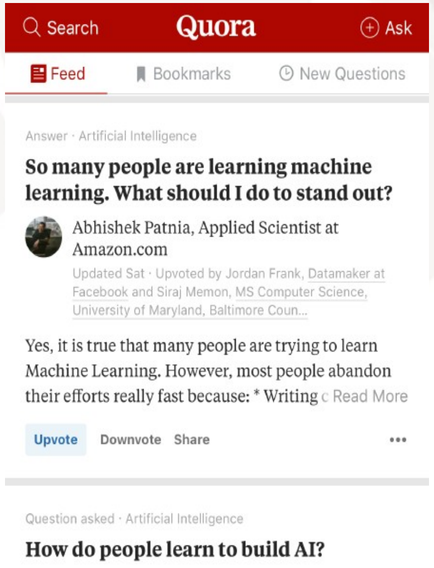
CONTENTS

Recommender System

- **Recommendation systems** help users discover items of interest from a large resource collection
- Recommender systems are everywhere, e.g., Amazon, Quora, Douban
- Recommender systems play a pivotal role in various online services







Collaborative filtering: a basic recommendation method

- Predict the interests of a user by collecting from many other users, e.g. matrix factorization
- Suffer from cold-start problem: data sparseness, new users/items

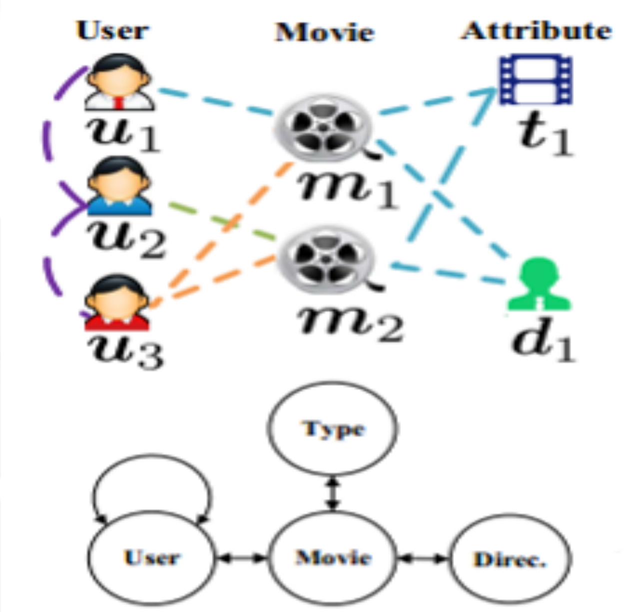
Integrate more auxiliary information

- Social network → social recommendation
- Location → location based recommendation
- Feature information → context based recommendation

Heterogeneous information network is a promising way to integrate auxiliary data.

Heterogeneous Information Network (HIN)

- Include multiple types of nodes or links
- Flexible to characterize heterogeneous data
- Contain rich semantics



Meta-path

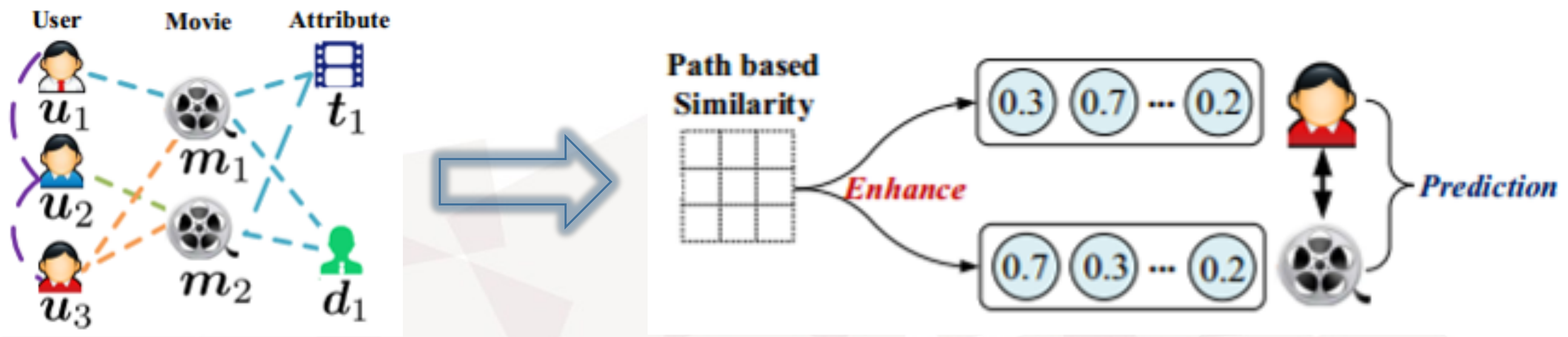
- A relation sequence connecting two objects in HIN
- Extract structural features
- Embody path semantics

$$User \xrightarrow{view} Movie \xrightarrow{viewed} User\ (UMU)$$

Two users view the same movies

$$User \xrightarrow{view} Movie \xrightarrow{directed} Director \xrightarrow{direct} Movie(UMDM)$$

Movies having the same type with the movies that the user viewed

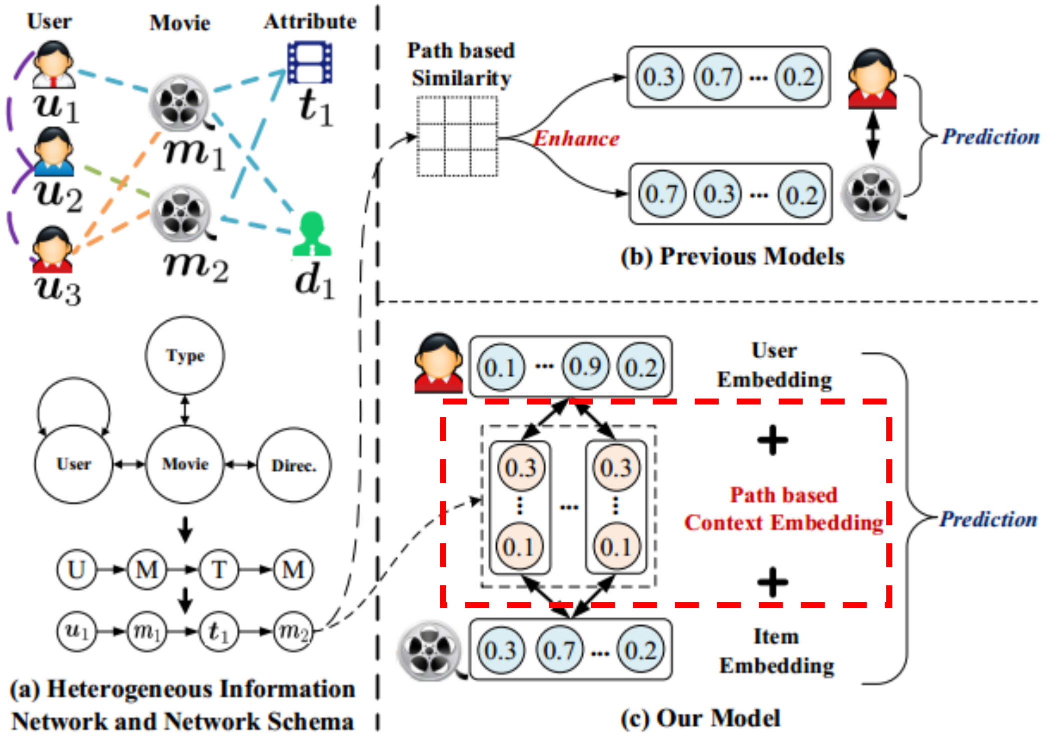


Existing HIN based methods

- Path based semantic relatedness as features for recommendation (e.g., OptRank, SemRec)
- Path based similarities for enhancing user/item representations (e.g., HeteRec, FMG)

Drawbacks

- Representations are not tailored for recommendation
- Seldom explicit representation for path/meta-path
- Only capture two way user-item interactions, without considering the mutual effect between user, item and path



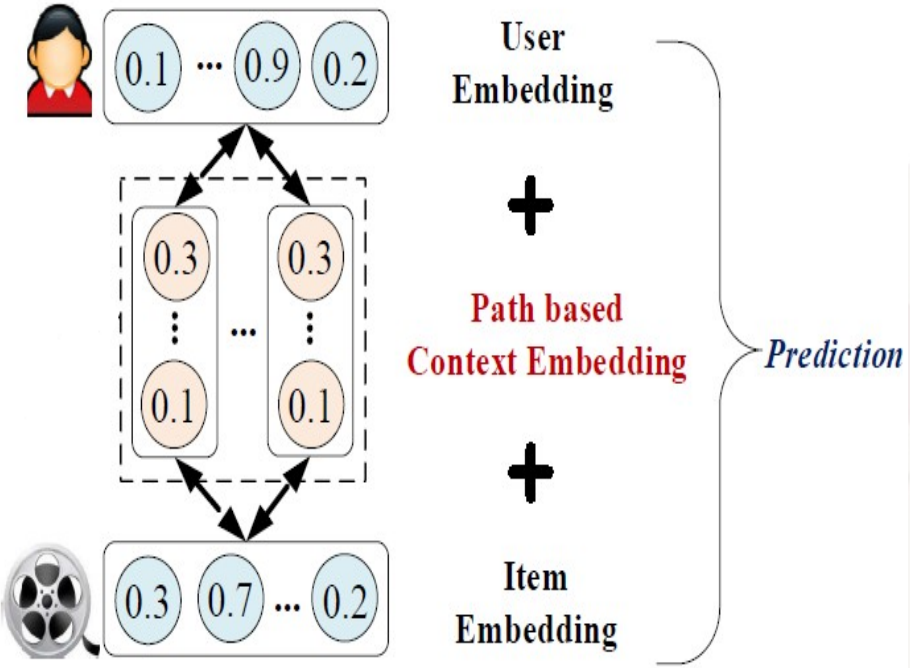
Our idea

- Learn explicit representations for meta-path based context tailored for the recommendation task
- Characterize a three-way interaction ⟨user, meta-path, item⟩

CONTENTS

KDD2018

Key Research Problems



Heterogeneity

Comprehensively and flexibly utilize heterogeneous information

Interpretability

Utilize context semantics for interpretable recommendation

Mutual Effect

Utilize the mutual effect between user-item pair and meta-path based context

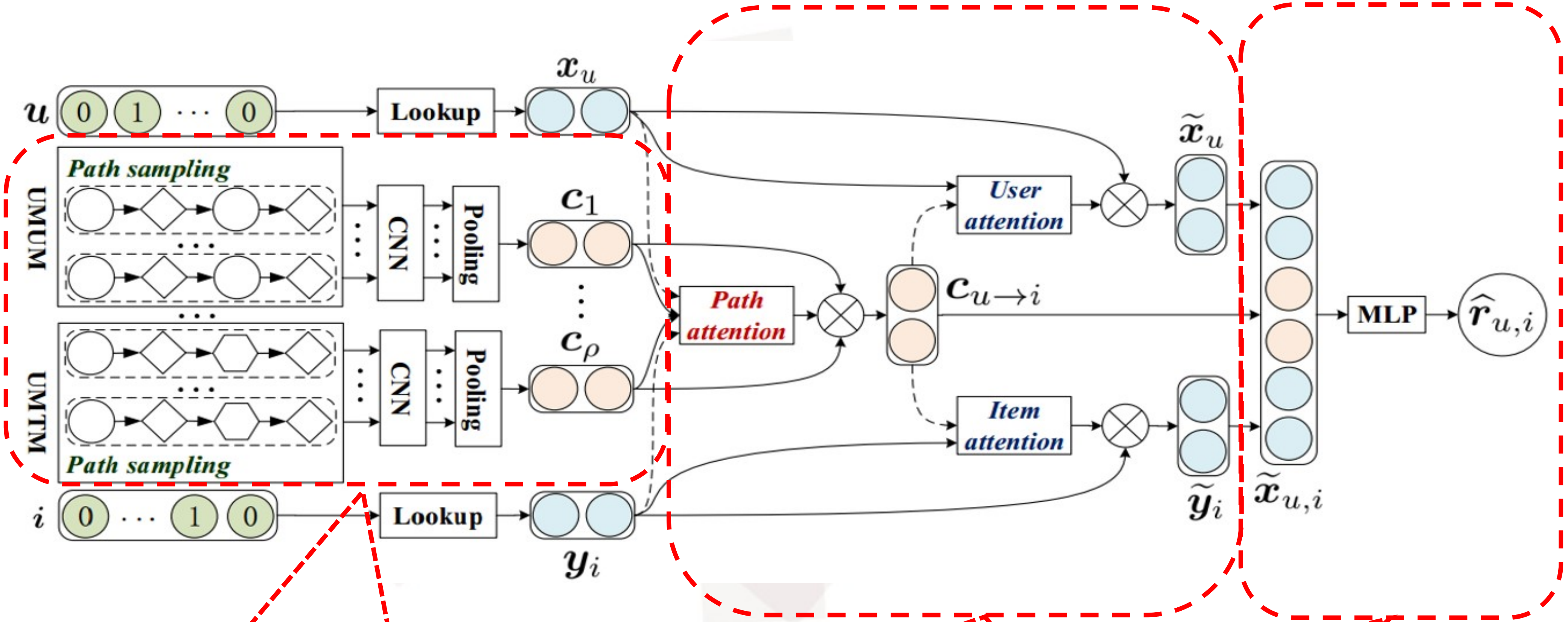
Rank

A more useful ranking model for HIN based recommendation

Meta-path based Context for RECommendation (MCRec)

Overall Framework

Heterogeneity → **A flexible deep NN based framework**

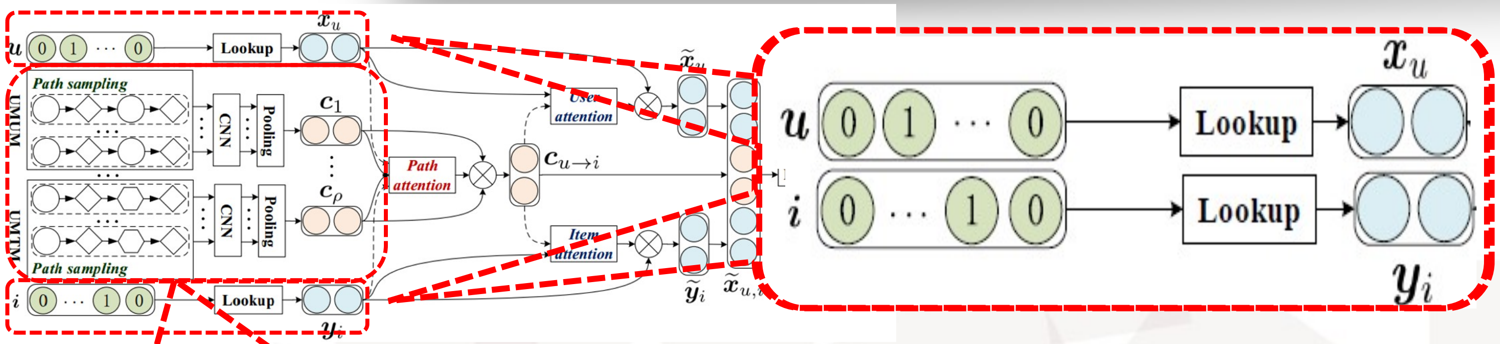


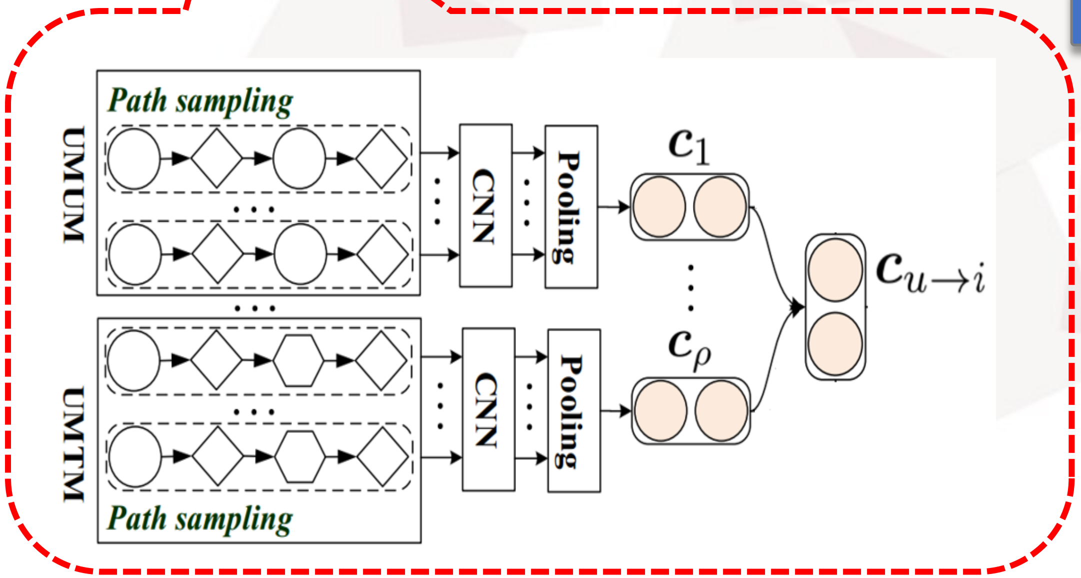
Interpretability → Meta-path based context embedding

Mutual Effect → Neural co-attention mechanism

Rank → Ranking predication model

Embedding Model for user, item and context





User/Item Embedding

- Look up

$$\mathbf{x}_u = \mathbf{P}^\top \cdot \mathbf{p}_u,$$
$$\mathbf{y}_i = \mathbf{Q}^\top \cdot \mathbf{q}_i.$$

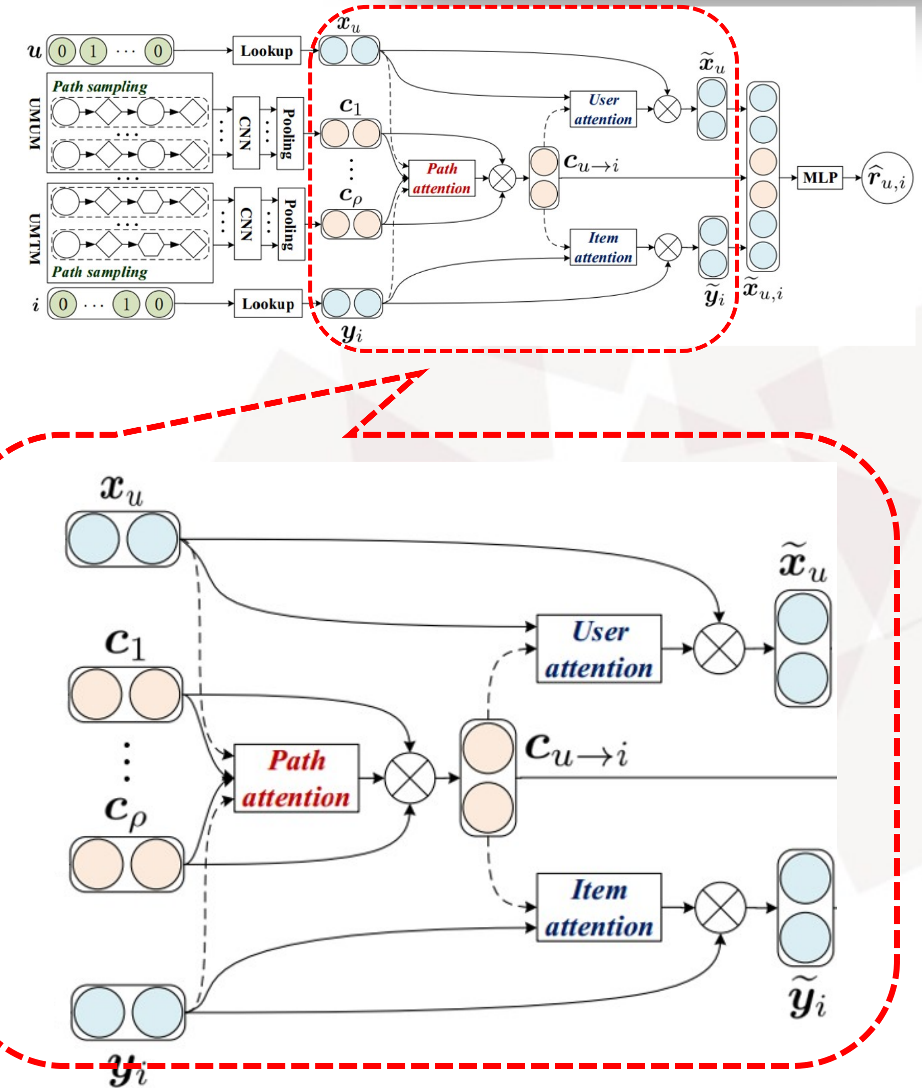
Meta-path based Context Embedding

- **Priority based Sampling Strategy**
  - SVD/FM for pre-training
  - Priority based random walk based on meta paths
- **CNN & Pooling for Encoding Context**

$$\mathbf{h}_p = CNN(\mathbf{X}^p; \Theta)$$
$$\mathbf{c}_\rho = \text{max-pooling}(\{\mathbf{h}_p\}_{p=1}^K).$$

- **Merge**

$$\mathbf{c}_{u \to i} = \frac{1}{|\mathcal{M}_{u \to i}|} \sum_{\rho \in \mathcal{M}_{u \to i}} \mathbf{c}_\rho.$$



Neural Co-attention Model

Path Attention Part

Attention score

$$\boldsymbol{\alpha}^{(1)}_{u,i,\rho} = f(\mathbf{W}^{(1)}_u \mathbf{x}_u + \mathbf{W}^{(1)}_i \mathbf{y}_i + \mathbf{W}^{(1)}_\rho \mathbf{c}_\rho + \mathbf{b}^{(1)}),$$

$$\alpha^{(2)}_{u,i,\rho} = f(\mathbf{w}^{(2)\top} \boldsymbol{\alpha}^{(1)}_{u,i,\rho} + b^{(2)}),$$

Softmax

$$\alpha_{u,i,\rho} = \frac{\exp(\alpha^{(2)}_{u,i,\rho})}{\sum_{\rho' \in M_{u\to i}} \exp(\alpha^{(2)}_{u,i,\rho'})}.$$

Apply

$$\mathbf{c}_{u\to i} = \sum_{\rho \in M_{u\to i}} \alpha_{u,i,\rho} \cdot \mathbf{c}_\rho,$$

User and Item Attention Part

Attention score

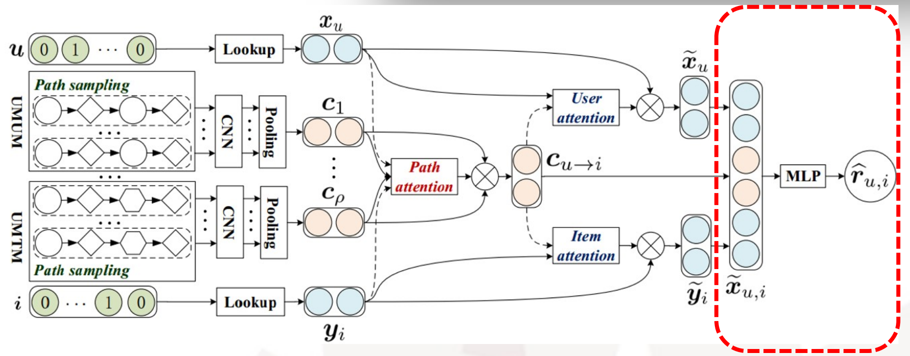
$$\boldsymbol{\beta}_u = f(\mathbf{W}_u \mathbf{x}_u + \mathbf{W}_{u\to i} \mathbf{c}_{u\to i} + \mathbf{b}_u),$$

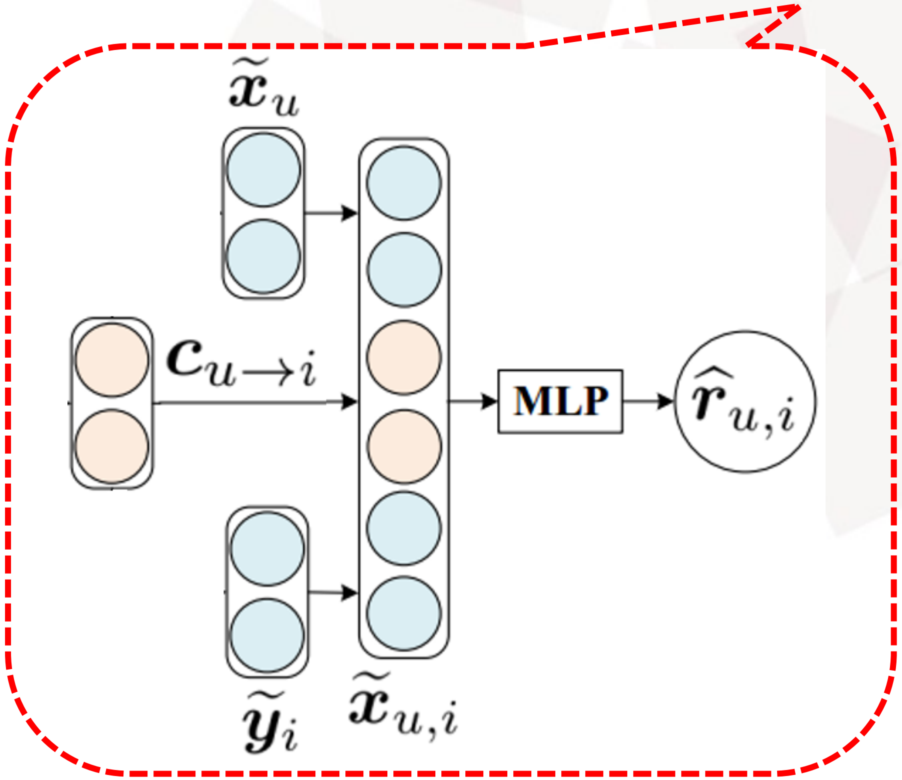
$$\boldsymbol{\beta}_i = f(\mathbf{W}'_i \mathbf{y}_i + \mathbf{W}'_{u\to i} \mathbf{c}_{u\to i} + \mathbf{b}'_i),$$

Apply

$$\tilde{\mathbf{x}}_u = \boldsymbol{\beta}_u \odot \mathbf{x}_u,$$

$$\tilde{\mathbf{y}}_i = \boldsymbol{\beta}_i \odot \mathbf{y}_i.$$





Ranking Prediction Model based on MLP

Concatenate

$$\widetilde{\mathbf{x}}_{u,i} = \tilde{\mathbf{x}}_u \oplus \mathbf{c}_{u\to i} \oplus \tilde{\mathbf{y}}_i,$$

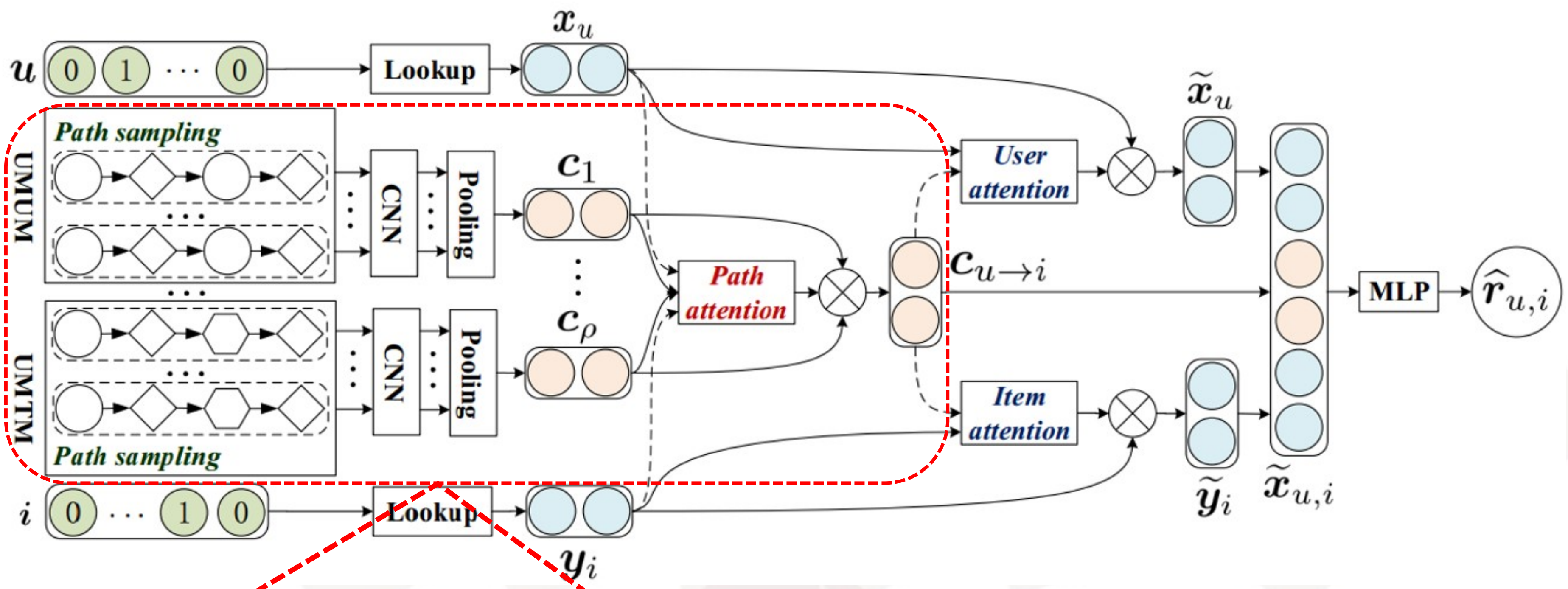
Multi-layer Perceptron

$$\widetilde{\mathbf{x}}_1 = f(\mathbf{W}_0 \widetilde{\mathbf{x}} + b_0)$$
$$\cdots\cdots$$
$$\widetilde{\mathbf{x}}_L = f(\mathbf{W}_{L-1} \widetilde{\mathbf{x}}_{L-1} + b_{L-1})$$
$$\hat{r}_{u,i} = \sigma(w^T \widetilde{\mathbf{x}}_L)$$

Optimization with negative sampling

$$\ell_{u,i} = -\log \hat{r}_{u,i} - E_{j\sim P_{neg}}[\log(1-\hat{r}_{u,j})],$$

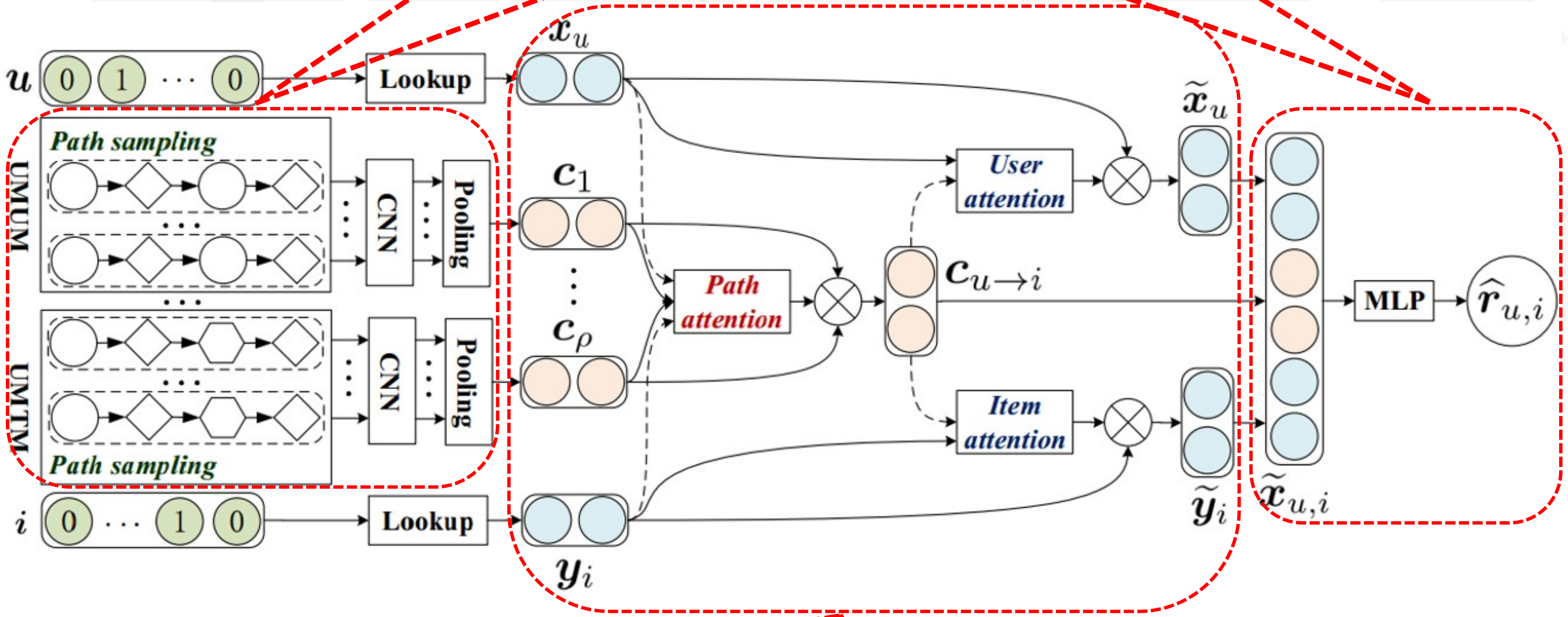
VS traditional recommendations



- Add explicit representation of meta-path based context
- Flexibly leverage heterogeneous information

VS HIN based recommendations

Deep feature exploration with NN



Exploit mutual effect between users, items and context

CONTENTS

Datasets

Datasets	Relations (A-B)	#A	#B	#A-B	Meta-paths
Movielens	*User-Movie*	943	1,682	100,000	UMUM
	User-User	943	943	47,150	UMGM
	Movie-Movie	1,682	1,682	82,798	UUUM
	Movie-Genre	1,682	18	2861	UMMM
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

Metrics

- Perc@10
- Recall@10
- NDCG@10

Baselines

CF based Methods

- ItemKNN
- BPR
- MF
- NeuMF

HIN based Methods

- SVDFeature_{hete}
- SVDFeature_{mp}
- HeteRS
- FMG_{rank}

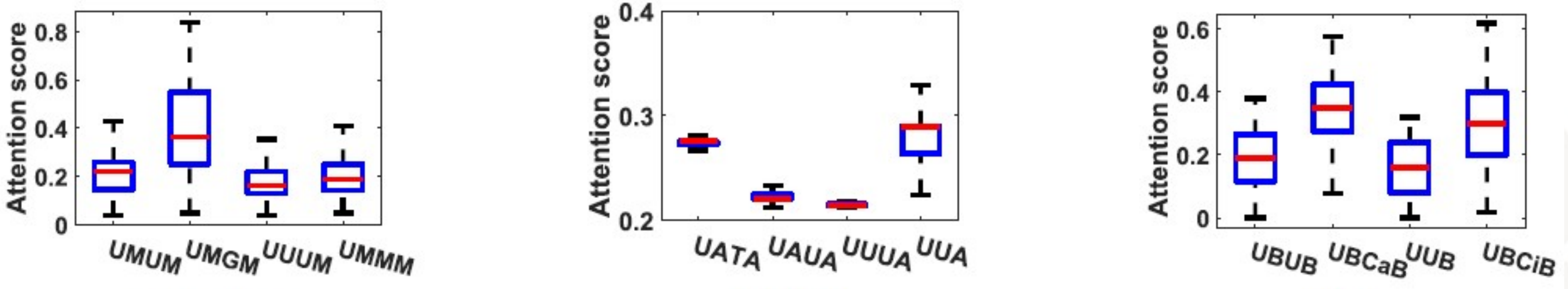
Our Methods

- MCRec_{rand}
- MCRec_{avg}
- MCRec_{mp}
- MCRec

Model	Movielens			LastFM			Yelp		
	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10
ItemKNN	0.2578	0.1536	0.5692	0.4160	0.4513	0.7981	0.1386	0.5421	0.5378
BRP	0.3010	0.1946	0.6459	0.4129	0.4492	0.8099	0.1474	0.5504	0.5549
MF	0.3247	0.2053	0.6511	0.4364	0.4634	0.7921	0.1503	0.5350	0.5322
NeuMF	0.3293*	0.2090	0.6587	0.4540	0.4678	0.8104	0.1504	0.5857	0.5713
SVDFeature$_{hete}$	0.3171	0.2021	0.6445	0.4576	0.4841	0.8290*	0.1404	0.5613	0.5289
SVDFeature$_{mp}$	0.3109	0.1929	0.6536	0.4391	0.4651	0.8116	0.1524	0.5932	0.5974*
HeteRS	0.2485	0.1674	0.5967	0.4276	0.4489	0.8026	0.1423	0.5613	0.5600
FMG$_{rank}$	0.3256	0.2165*	0.6682*	0.4630*	0.4916*	0.8263	0.1538*	0.5951*	0.5861
MCRec$_{rand}$	0.3223	0.2104	0.6650	0.4540	0.4795	0.8002	0.1510	0.5842	0.5718
MCRec$_{avg}$	0.3270	0.2111	0.6631	0.4645	0.4914	0.8311	0.1595	0.5933	0.6021
MCRec$_{mp}$	0.3401	0.2200	0.6828	0.4662	0.4924	0.8428	0.1655	0.6303	0.6228
MCRec	**0.3451**#	**0.2256**#	**0.6900**#	**0.4807**#	**0.5068**#	**0.8526**#	**0.1686**#	**0.6326**#	**0.6301**#

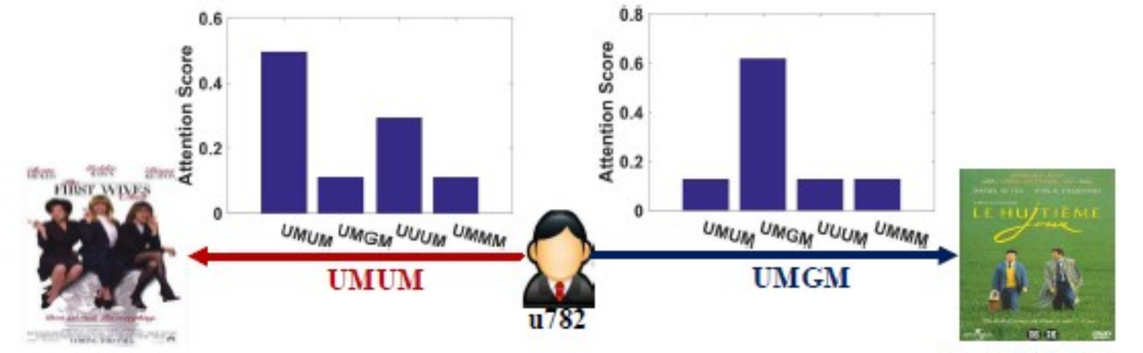
MCRec significantly outperforms CF, NN, and HIN based recommendations

Distribution of attention weights



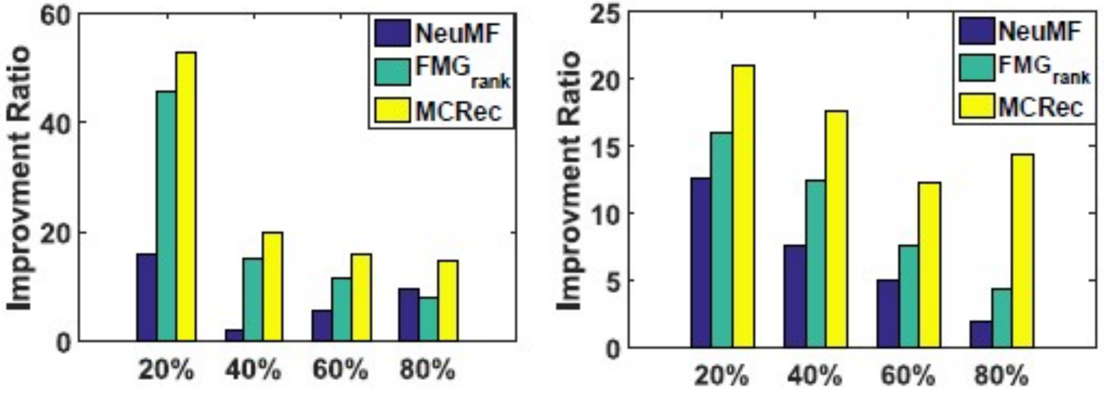
(a) Movielens (b) LastFM (c) Yelp

Case study on Movielens dataset



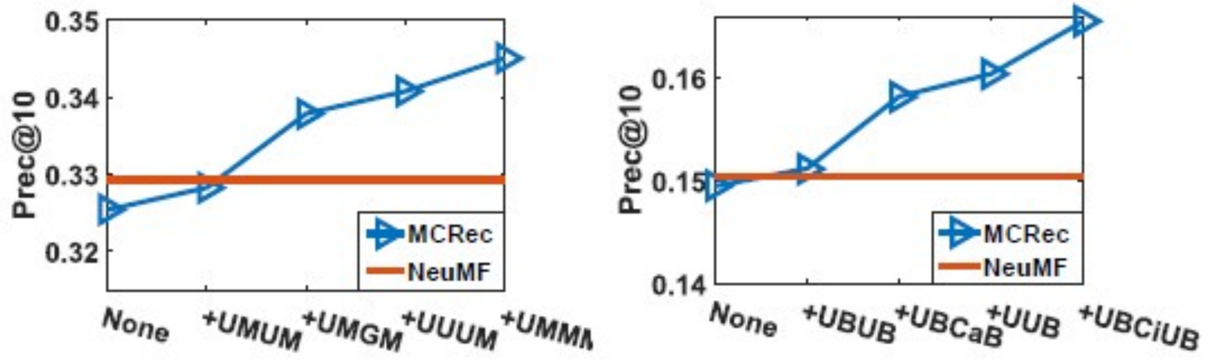
MCRec provides personalized interpretable recommendation

Cold-start recommendation



(a) Movielens (b) Yelp

Impact of different meta-paths



(a) Movielens (b) Yelp

MCRec is promising for cold-start problem

CONTENTS

KDD2018

- We designed a three-way neural interaction model by explicitly incorporating meta-path based context

- The co-attention model mutually improved the representations for path based context, users and items

- Extensive experimental results have revealed the effectiveness and interpretability of our model

KDD2018

Thanks
Q&A

北京邮電大學 BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

中国人民大學 RENMIN UNIVERSITY OF CHINA

UNIVERSITY OF ILLINOIS — LEARNING & LABOR — CHARTERED 1867

More materials in www.shichuan.org