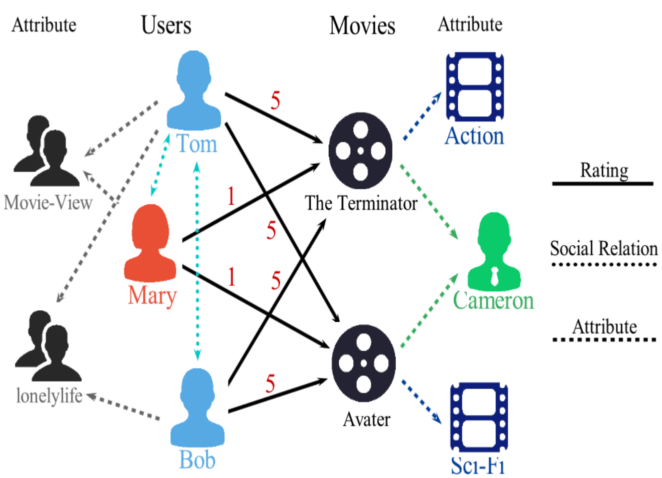
北京邮電大學
BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

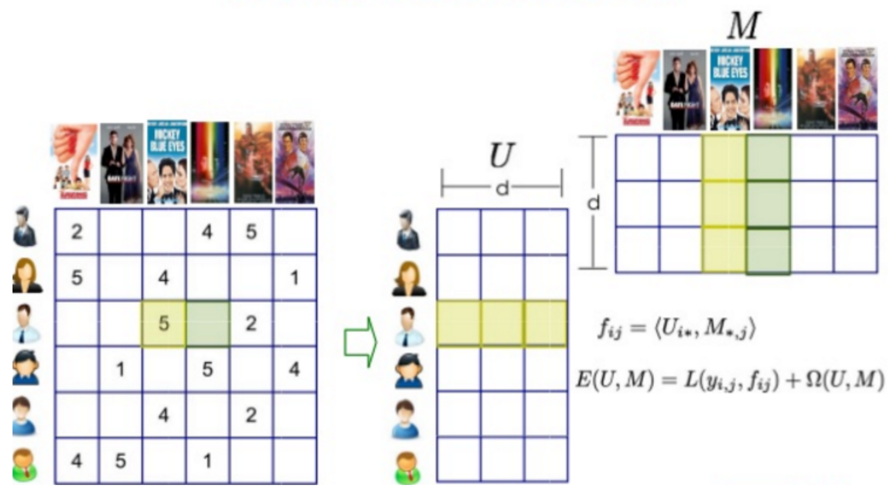
TSEG
Telecommunications Software Engineering Group

Heterogeneous Information Network Embedding for Recommendation

通信软件工程中心

Trandition Recommendation

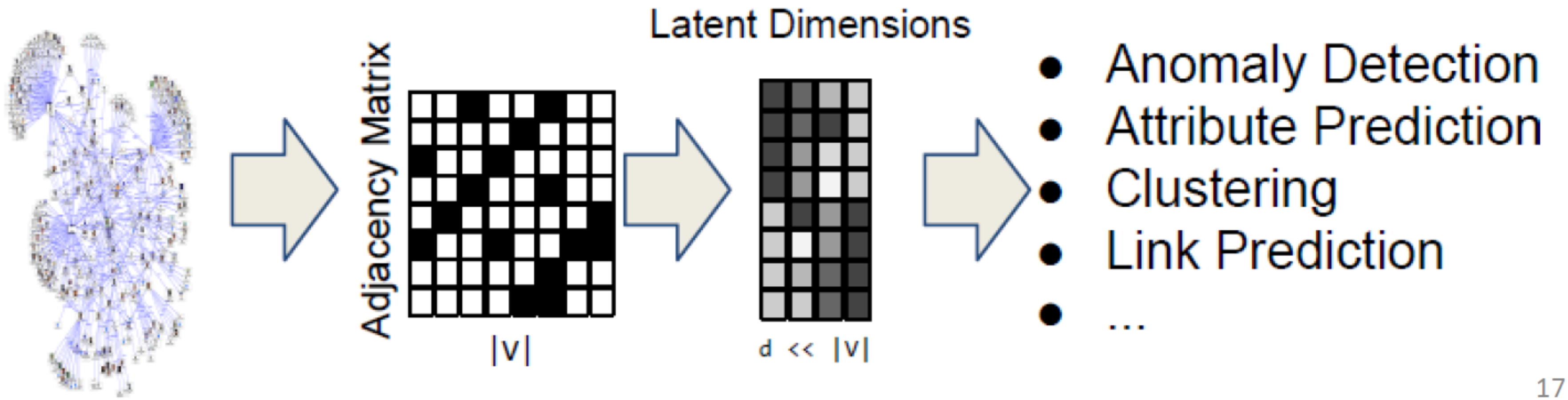




Matrix Factorization —weakness→ **Cold-start problem**

HIN-based Recommendation → **Limited ability of information extraction and exploitation**

Network Embedding



Network Embedding

- Deepwalk
- Node2vec
- LINE
- SDNE
- …

weakness →

- Towards homogeneous network.
- Limited ability to uncover complex semantics of HIN

Framework

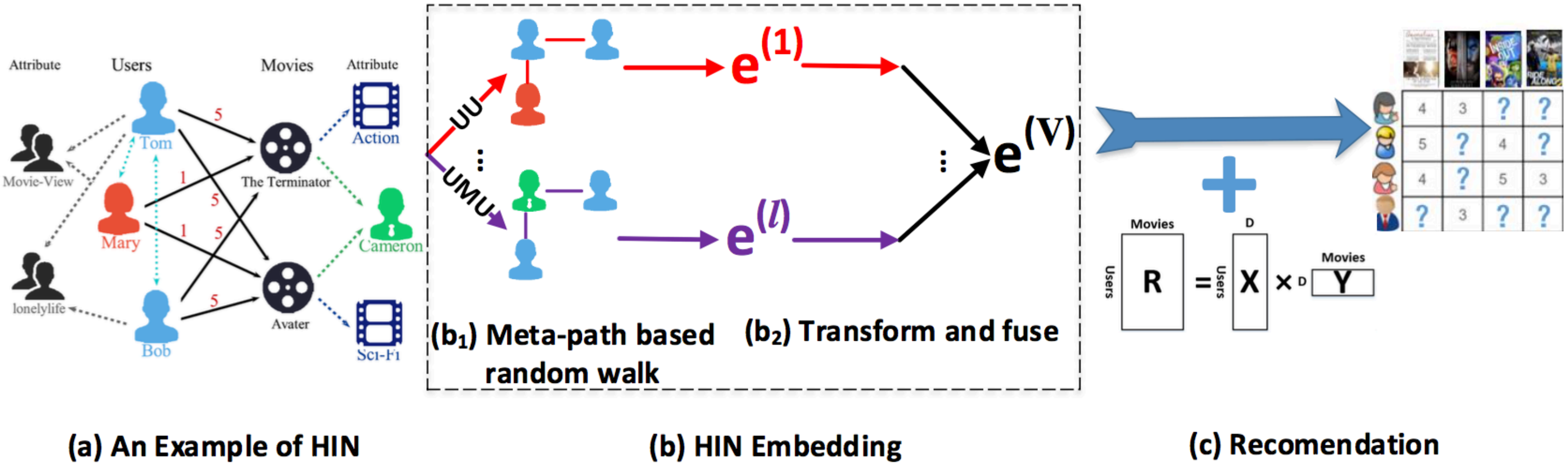


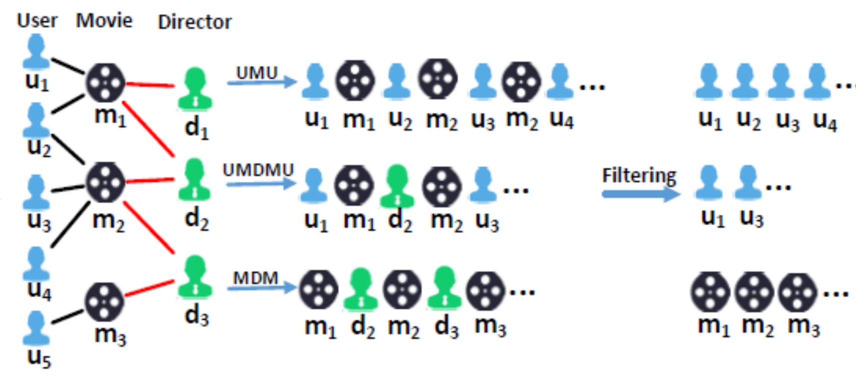
Figure 1: The schematic illustration of the proposed HERec approach.

HERec
- Heterogeneous Information Network Embedding
- Fuse Embeddings into MF for Recommendation

HIN Embedding

Metepath based random walk

Giving a heterogeneous network $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ and a meta-path $\rho : A_1 \xrightarrow{R_1} \cdots A_t \xrightarrow{R_t} A_{t+1} \cdots \xrightarrow{R_l} A_{l+1}$, the walk path is generated according to the following distribution:



$$P(n_{t+1} = x | n_t = v, \rho)$$

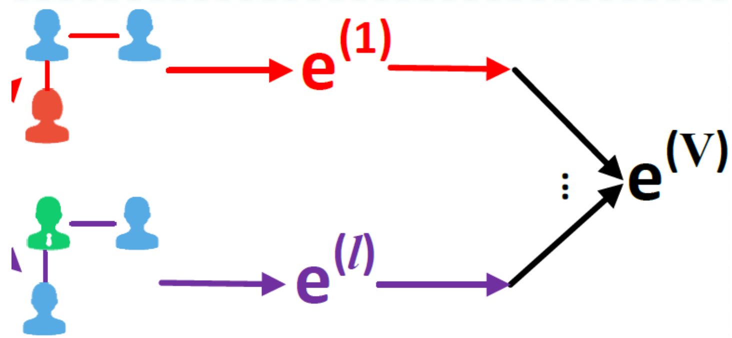
$$= \begin{cases} \frac{1}{|\mathcal{N}^{A_{t+1}}(v)|}, & (v, x) \in \mathcal{E} \text{ and } \phi(x) = A_{t+1}; \\ 0, & \text{otherwise}, \end{cases}$$

Type Constraints and Filtering

- Uncover heterogeneous information with homogeneous NE objective
- Utilize more neighbor information with a fix-length window

HIN Embedding

Embedding for single metapath



$$\max_{f} \sum_{u \in \mathcal{V}} \log Pr(\mathcal{N}_u | f(u)),$$

where $f : \mathcal{V} \rightarrow \mathbb{R}^d$ is a function (aiming to learn) maps each node to d-dimensional feature space, and $\mathcal{N}_u \subset \mathcal{V}$ represents the neighborhood of node u, *w.r.t.* to a specific meta-path. We can learn the embedding mapping function $f(\cdot)$ by applying stochastic gradient descent (SGD) to optimize

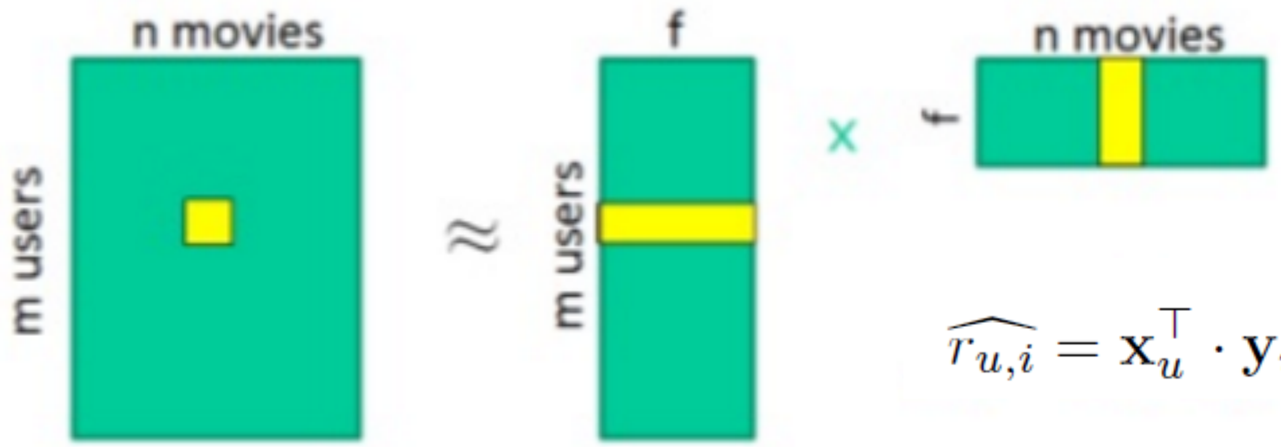
Embedding Fusion

- Uncover different information of various meta-paths
- A good fusion function should be learned according to the specific task

$$\begin{aligned} \mathbf{e}_u^{(U)} &\leftarrow g(\{\mathbf{e}_u^{(l)}\}), \\ \mathbf{e}_i^{(I)} &\leftarrow g(\{\mathbf{e}_i^{(l)}\}), \end{aligned}$$

Recommendation

Based Rating Prediction



$$\widehat{r_{u,i}} = \mathbf{x}_u^\top \cdot \mathbf{y}_i,$$

Extended Rating Prediction

- Integrated with HIN Embeddings

$$\widehat{r_{u,i}} = \mathbf{x}_u^\top \cdot \mathbf{y}_i + \alpha \cdot \boldsymbol{e}_u^{(U)\top} \cdot \boldsymbol{\gamma}_i^{(I)} + \beta \cdot \boldsymbol{\gamma}_u^{(U)\top} \cdot \boldsymbol{e}_i^{(I)}, \quad (5)$$

Recommendation

Fusion functions

- Simple linear fusion. We use a global set of meta-path transformation matrices to combine user embeddings

$$g(\{\boldsymbol{e}_u^{(l)}\}) = \frac{1}{|\mathcal{P}|}\sum_{l=1}^{|\mathcal{P}|}(\mathbf{M}^{(l)}\boldsymbol{e}_u^{(l)} + \boldsymbol{b}^{(l)}), \qquad (6)$$

- Personalized linear fusion. The simple linear fusion cannot model users' personalized preference over the meta-paths. So we further incorporate user-specific meta-path weights

$$g(\{\boldsymbol{e}_u^{(l)}\}) = \frac{1}{|\mathcal{P}|}\sum_{l=1}^{|\mathcal{P}|} w_u^{(l)}(\mathbf{M}^{(l)}\boldsymbol{e}_u^{(l)} + \boldsymbol{b}^{(l)}), \qquad (7)$$

- Personalized non-linear fusion. Linear fusion has limited expressive power in modeling complex data relations. Hence, we use non-linear function to enhance the fusion ability

$$g(\{\boldsymbol{e}_u^{(l)}\}) = \sigma\left(\sum_{l=1}^{|\mathcal{P}|} w_u^{(l)}\sigma\left(\mathbf{M}^{(l)}\boldsymbol{e}_u^{(l)} + \boldsymbol{b}^{(l)}\right)\right), \qquad (8)$$

Recommendation

Model Objective

$$\mathcal{L} = \sum_{\langle u,i,r_{u,i}\rangle \in \mathcal{R}} (r_{u,i} - \widehat{r_{u,i}})^2 + \lambda \sum_{u} (\|\mathbf{x}_u\|_2 + \|\mathbf{y}_i\|_2$$
$$+ \|\boldsymbol{\gamma}_u^{(U)}\|_2 + \|\boldsymbol{\gamma}_i^{(I)}\|_2 + \|\boldsymbol{\Theta}^{(U)}\|_2 + \|\boldsymbol{\Theta}^{(I)}\|_2), \qquad (9)$$

Model Learning

$$\boldsymbol{\Theta}_{u,l}^{(U)} \leftarrow \boldsymbol{\Theta}_{u,l}^{(U)} - \eta \cdot (-\alpha(r_{u,i} - \widehat{r_{u,i}})\boldsymbol{\gamma}_i^{(I)} \frac{\partial \mathbf{e}_u^{(U)}}{\boldsymbol{\Theta}_{u,l}^{(U)}} + \lambda_\Theta \boldsymbol{\Theta}_{u,l}^{(U)}),$$

$$\boldsymbol{\gamma}_u^{(U)} \leftarrow \boldsymbol{\gamma}_u^{(U)} - \eta \cdot (-\beta(r_{u,i} - \widehat{r_{u,i}})\mathbf{e}_i^{(I)} + \lambda_\gamma \boldsymbol{\gamma}_u^{(U)}).$$

We omit the learning of $\boldsymbol{\Theta}_{i,l}^{(I)}$ and $\boldsymbol{\gamma}_i^{(I)}$ due to the similar derivations.

Dataset

Table 1: Statistics of the three datasets.

Dataset (Density)	Relations (A-B)	Number of A	Number of B	Number of (A-B)
Douban Movie (0.63%)	User-Movie	13,367	12,677	1,068,278
	User-User	2,440	2,294	4,085
	User-Group	13,337	2,753	570,047
	Movie-Director	10,179	2,449	11,276
	Movie-Actor	11,718	6,311	33,587
	Movie-Type	12,678	38	27,668
Douban Book (0.27%)	User-Book	13,024	22,347	792,026
	User-User	12,748	12,748	169,150
	Book-Author	21,907	10,805	21,905
	Book-Publisher	21,773	1,815	21,773
	Book-Year	21,192	64	21,192
Yelp (0.08%)	User-Business	16,239	14,284	198,397
	User-User	10,580	10,580	158,590
	User-Compliment	14,411	11	76,875
	Business-City	14,267	47	14,267
	Business-Category	14,180	511	40,009

Metric and Settings

Metric

- $RMSE = \frac{\sqrt{\sum_{(u,i)\in R}(r_{u,i}-\hat{r}_{u,i})^2}}{|R|}$
- $MAE = \frac{\sum_{(u,i)\in R}|r_{u,i}-\hat{r}_{u,i}|}{|R|}$

Table 2: The selected meta-paths used in our work.

Dataset	Meta-paths
Douban Movie	UMU, UMDMU, UMAMU, UMTMU MUM, MAM, MDM, MTM
Douban Book	UBU, UBABU, UBPBU, UBYBU BUB, BPB, BYB
Yelp	UBU, UBCiBU, UBCaBU BUB, BCiB, BCaB

Compared Methods

Typical methods
- PMF
- SoMF

NE-based methods
- HERec_{dw}
- HERec_{mp}

HIN-based Methods
- FM_{HIN}
- HeteMF
- SemRec
- DSR

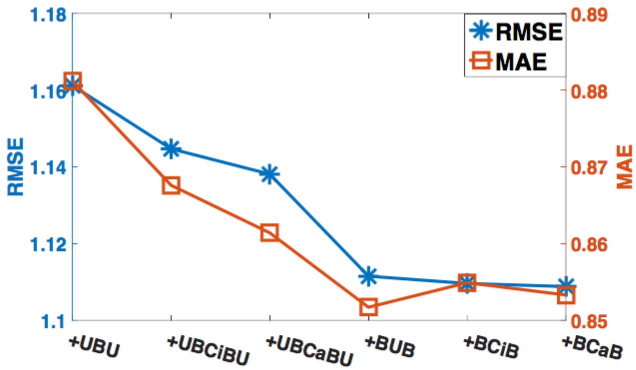
Variants of HERec
- HERec_{sl}
- HERec_{pl}
- HERec_{pnl}

Effectiveness Experiments

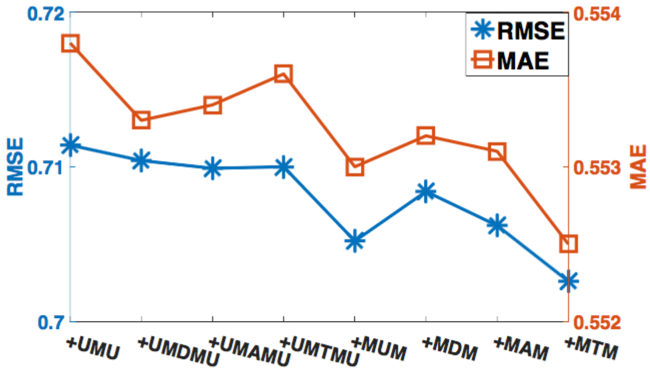
Table 3: Results of effectiveness experiments on three datasets. A smaller value indicates a better performance.

Dataset	Training	Metrics	PMF	SoMF	FM_{HIN}	HeteMF	SemRec	DSR	$HERec_{dw}$	$HERec_{mp}$	$HERec_{sl}$	$HERec_{pl}$	$HERec_{pnl}$
Douban Movie	80%	MAE	0.5741	0.5817	0.5696	0.5750	0.5695	0.5681	0.5703	0.5515	0.5617	0.5523	**0.5519**
		RMSE	0.7641	0.7680	0.7248	0.7556	0.7399	0.7225	0.7446	0.7121	0.7216	**0.7024**	0.7053
	60%	MAE	0.5867	0.5991	0.5769	0.5894	0.5738	0.5831	0.5838	0.5611	0.5711	0.5606	**0.5587**
		RMSE	0.7891	0.7950	0.7342	0.7785	0.7551	0.7408	0.7670	0.7264	0.7336	**0.7142**	0.7148
	40%	MAE	0.6078	0.6328	0.5871	0.6165	0.5945	0.6170	0.6073	0.5747	0.5832	0.5732	**0.5699**
		RMSE	0.8321	0.8479	0.7563	0.8221	0.7836	0.7850	0.8057	0.7429	0.7514	0.7334	**0.7315**
	20%	MAE	0.7247	0.6979	0.6080	0.6896	0.6392	0.6584	0.6699	0.6063	0.5953	0.5965	**0.5900**
		RMSE	0.9440	0.9852	0.7878	0.9357	0.8599	0.8345	0.9076	0.7877	0.7916	0.7674	**0.7660**
Douban Book	80%	MAE	0.5774	0.5756	0.5716	0.5740	0.5675	0.5740	0.5875	0.5591	0.5578	0.5556	**0.5502**
		RMSE	0.7414	0.7302	0.7199	0.7360	0.7283	0.7206	0.7450	0.7081	0.7079	0.7093	**0.6811**
	60%	MAE	0.6065	0.5603	0.5812	0.5823	0.5833	0.6020	0.6203	0.5666	0.5690	0.5669	**0.5600**
		RMSE	0.7908	0.7518	0.7319	0.7466	0.7505	0.7552	0.7905	0.7318	0.7251	0.7274	**0.7123**
	40%	MAE	0.6800	0.6161	0.6028	0.5982	0.6025	0.6271	0.6976	0.5954	0.5838	**0.5638**	0.5774
		RMSE	0.9203	0.7936	0.7617	0.7779	0.7751	0.7730	0.9022	0.7703	0.7490	0.7549	**0.7400**
	20%	MAE	1.0344	0.6327	0.6396	0.6311	0.6481	0.6300	1.0166	0.6785	**0.6232**	0.6347	0.6450
		RMSE	1.4414	0.8236	0.8188	0.8304	0.8350	0.8200	1.3205	0.8869	**0.8168**	0.8382	0.8581
Yelp	90%	MAE	1.0412	1.0095	0.9013	0.9487	0.9043	0.9054	1.0388	0.8822	0.8643	0.8506	**0.8395**
		RMSE	1.4268	1.3392	1.1417	1.2549	1.1637	1.1186	1.3581	1.1309	1.1204	1.0948	**1.0907**
	80%	MAE	1.0791	1.0373	0.9038	0.9654	0.9176	0.9098	1.0750	0.8953	0.8789	0.8578	**0.8475**
		RMSE	1.4816	1.3782	1.1497	1.2799	1.1771	1.1208	1.4075	1.1516	1.1403	1.1139	**1.1117**
	70%	MAE	1.1170	1.0694	0.9108	0.9975	0.9407	0.9429	1.1196	0.9043	0.8889	0.8650	**0.8580**
		RMSE	1.5387	1.4201	1.1651	1.3229	1.2108	1.1582	1.4632	1.1639	1.1599	**1.1229**	1.1256
	60%	MAE	1.1778	1.1135	0.9435	1.0368	0.9637	1.0043	1.1691	0.9257	0.9042	0.8818	**0.8759**
		RMSE	1.6167	1.4748	1.2039	1.3713	1.2380	1.2257	1.5182	1.1887	1.1817	**1.1451**	1.1488
Average Rank			10.33	8.79	5.25	7.83	6.54	6.13	9.5	4.17	3.2	2.4	**1.79**

Detailed Analysis



(a) Yelp



(b) Douban Movie

Figure 4: Performance change of HERec when gradually incorporating meta-paths.

Detailed Analysis

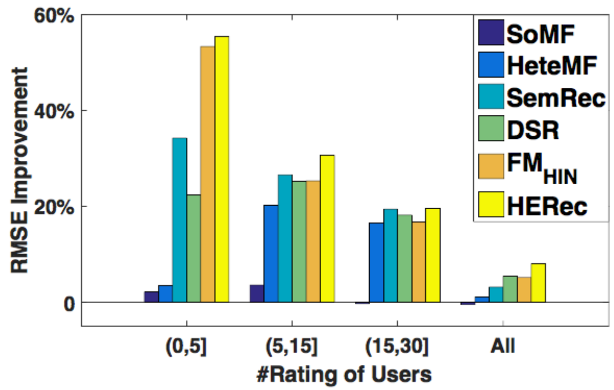


Figure 3: Performance comparison of different methods for cold-start prediction on Douban Movie dataset. y-axis denotes the improvement ratio over PMF.

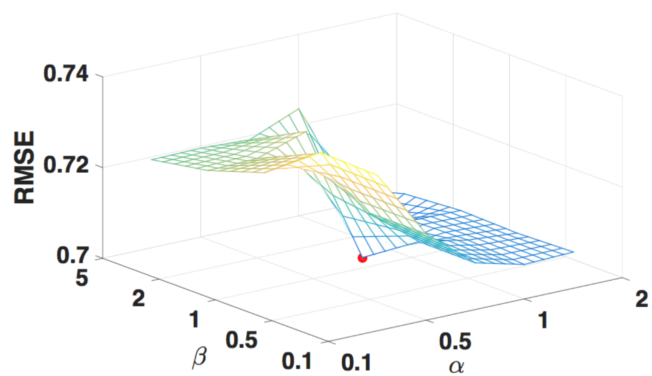


Figure 5: Varying parameters α and β on Douban Movie dataset.

谢谢