Cross-Domain Collaborative Filtering

Shuang Hong Yang

CSE, Gerogia Tech

Aug. 11th, 2010

Outline

- Problem Definition
- Different scenarios
- Representative work
- Summary and Discussion

Problem Definition

- Use movie ratings to help book recommendation

Movie

user\item	Harry Potter	God Farther	Avatar
Joe	3	5	5
Harry	?	4	5
George	5	3	3

Book

user\item	Harry Potter	God Farther	1984
Joe	3	?	?
Harry	?	?	5
David	4	?	5

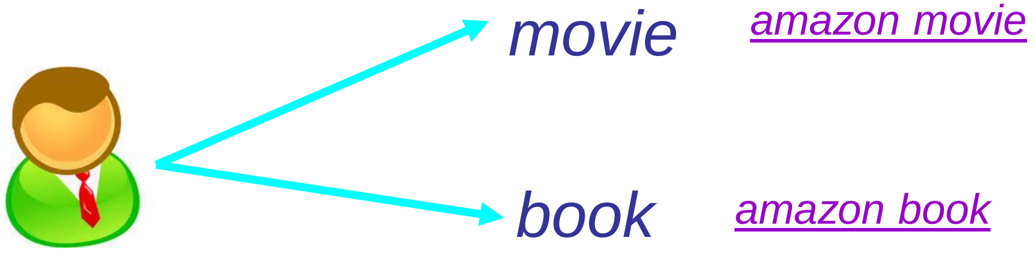
Why it is important?

- Behavior Integration: user generated data are scattered among different domains (systems, websites, categories)

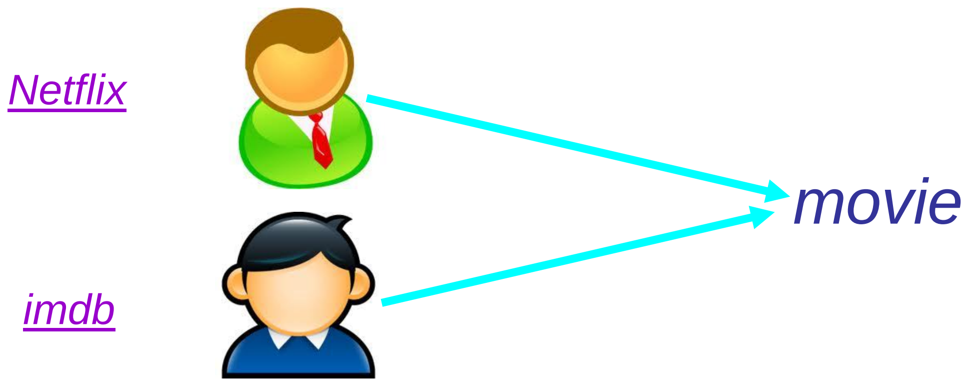
- New system: CF is effective for relatively mature systems, but not for low-traffic ones.

Typical settings

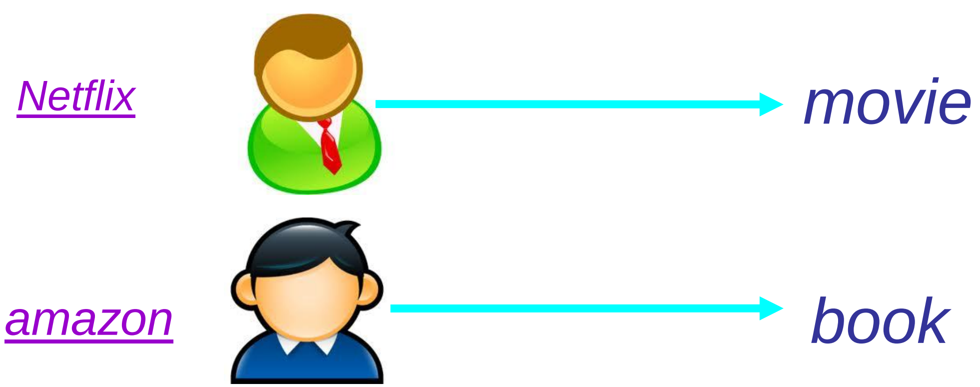
- Shared users



- Shared items



- Isolated



Typical settings

- Shared users : rating matrix

$$R = \begin{array}{c} \text{user}\left\{ \begin{bmatrix} R_m & \vdots & R_b \end{bmatrix} \right. \\ \qquad \text{movie} \quad \text{book} \end{array}$$

- Shared items

$$R = \begin{bmatrix} R_1 \\ \dots \\ R_2 \end{bmatrix}$$

- Isolated

$$R = \begin{bmatrix} R_1 & \\ & R_2 \end{bmatrix}$$

Collaborative Filtering

- Single domain

$$\boxed{\begin{aligned} R &= UV^\top + (e_{uv}) \\ U_u &\sim \mathcal{N}(0, \sigma^2 I) \\ V_v &\sim \mathcal{N}(0, \sigma^2 I) \\ e &\sim \mathcal{N}(0, \sigma^2) \end{aligned}}$$

Transfer Collaborative Filtering

- Shared users
  - shared factor [c.f. collaborative matrix factorization]

$$R = \begin{bmatrix} R_1 & R_2 \end{bmatrix} = \begin{bmatrix} UV_1 & UV_2 \end{bmatrix}$$

  - correlated factor

$$R = \begin{bmatrix} R_1 & R_2 \end{bmatrix} = \begin{bmatrix} U_1V_1 & U_2V_2 \end{bmatrix}$$

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \sim \mathcal{N}(0, \begin{bmatrix} \sigma_1^2 I & \Sigma \\ \Sigma^\top & \sigma_2^2 I \end{bmatrix})$$

Transfer Collaborative Filtering

- Shared items
  - shared factor [c.f. collaborative matrix factorization]

$$R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} U_1 V \\ U_2 V \end{bmatrix}$$

  - correlated factor

$$R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} U_1 V_1 \\ U_2 V_2 \end{bmatrix}$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \sim \mathcal{N}(0, \begin{bmatrix} \sigma_1^2 I & \Sigma \\ \Sigma^\top & \sigma_2^2 I \end{bmatrix})$$

Transfer Collaborative Filtering

- Isolated case

different users make decisions (rate) regarding different items

$$R = \begin{bmatrix} R_1 & \\ & R_2 \end{bmatrix} = \begin{bmatrix} U_1 V_1 & \\ & U_2 V_2 \end{bmatrix}$$

- shared factor does not work --- no correspondence to share
- correlated factor does not work either --- trivial covariance
- nothing useful to transfer?
  - Behavioral pattern regardless user/item
  - ...

Outline

- Problem Definition
- Different scenarios
- Representative work
- Summary and Discussion

Representative work

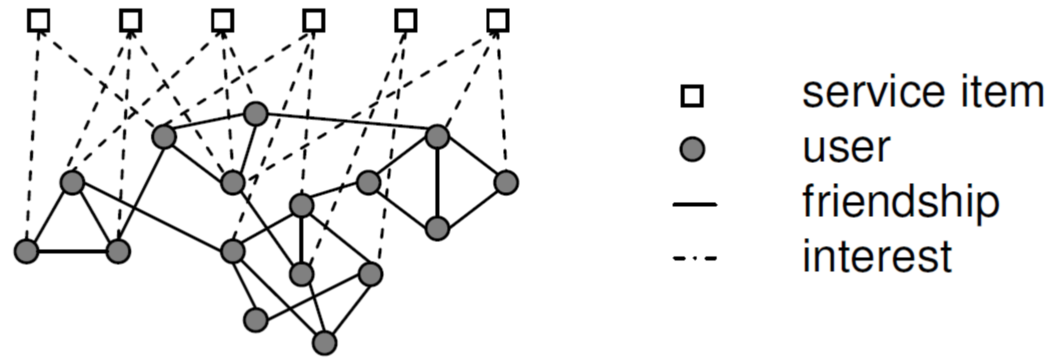
- Shared case
  - [Yang et al, WWW' 2011]
    Like like alike -- Joint friendship and interest propagation in social networks
  - [Zhang & Yeung, UAI' 2010]
    Multi-Domain Collaborative Filtering
  - [Pan et al, AAAI' 2010]: Transfer Learning in Collaborative Filtering for Sparsity Reduction
- Isolated case
  - [Li et al, IJCAI' 2009]: Can movies and books collaborate? - cross domain collaborative filtering for sparsity reduction
  - [Li et al, ICML' 2009]: Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model

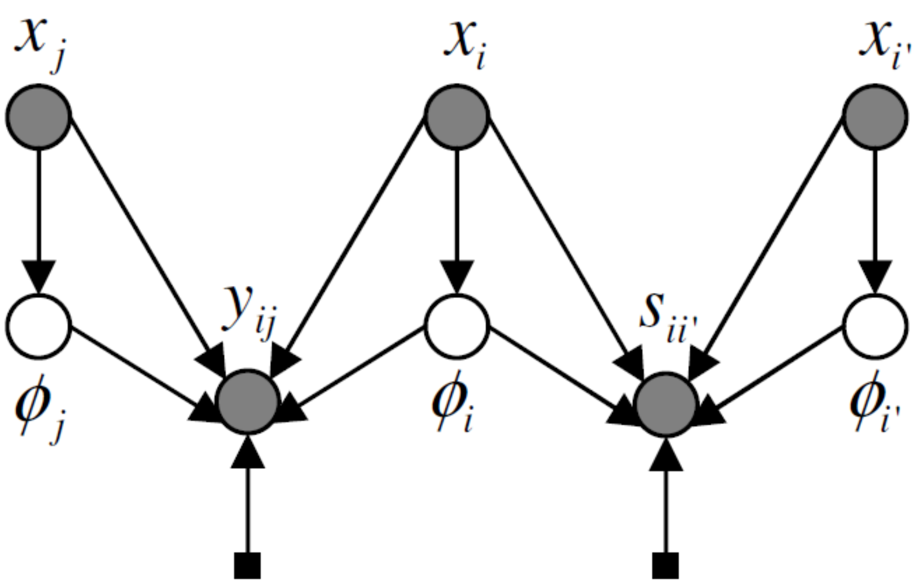
Representative work

- Shared case
  - [Yang et al, WWW' 2011]
    Like like alike -- Joint friendship and interest propagation in social networks
  - [Zhang & Yeung, UAI' 2010]
    Multi-Domain Collaborative Filtering
  - [Pan et al, AAAI' 2010]: Transfer Learning in Collaborative Filtering for Sparsity Reduction
- Isolated case
  - [Li et al, IJCAI' 2009]: Can movies and books collaborate? - cross domain collaborative filtering for sparsity reduction
  - [Li et al, ICML' 2009]: Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model

Like like alike ... [Yang et al *WWW' 2011*]

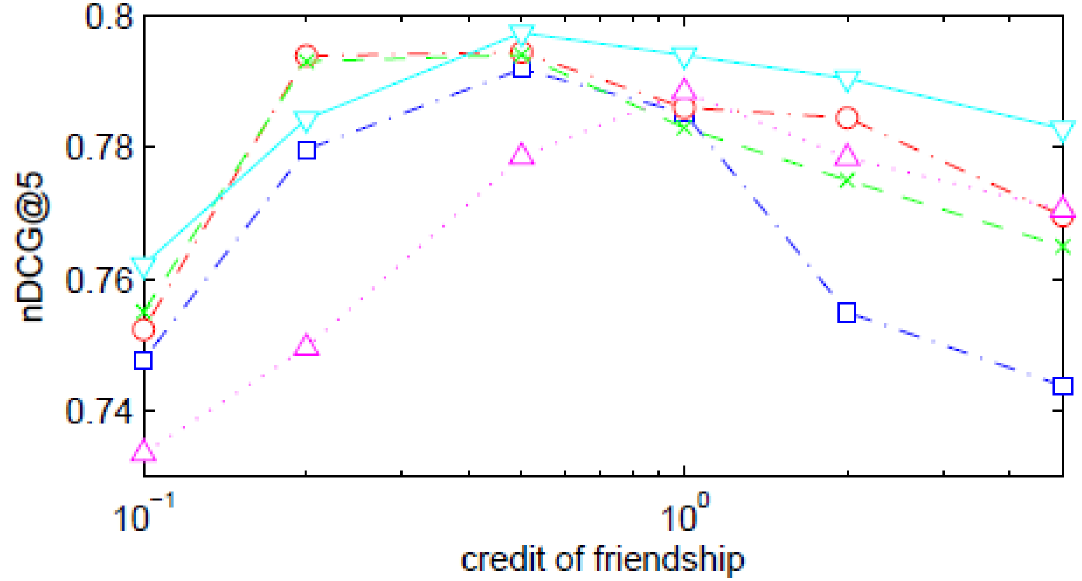
- Shared factor

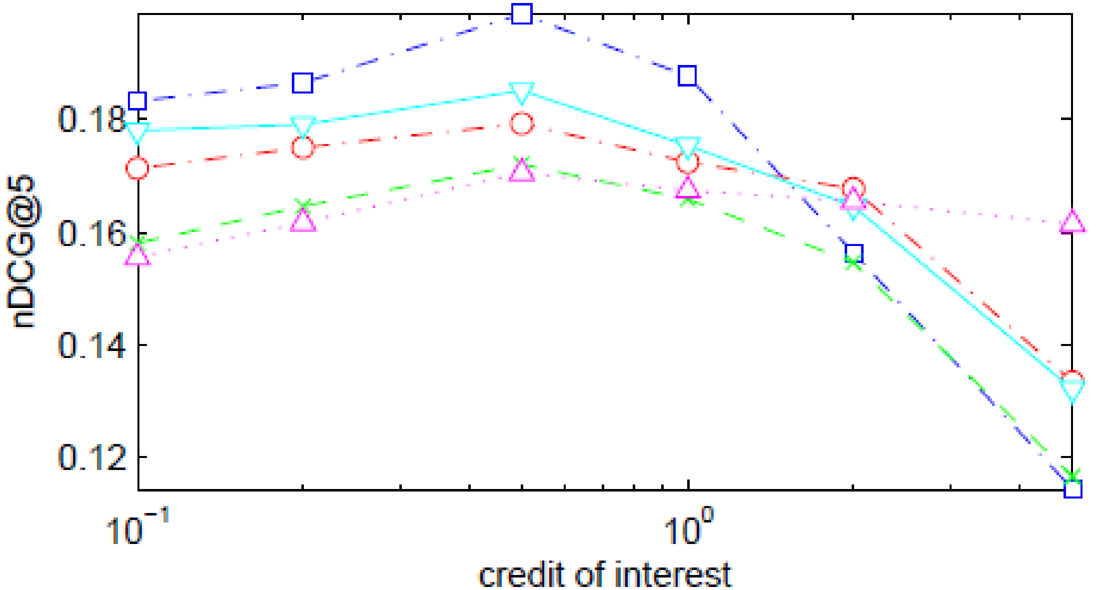




Like like alike ... [Yang et al *WWW' 2011*]

- Results





Representative work

- Shared case
  - [Yang et al, WWW' 2011]
    Like like alike -- Joint friendship and interest propagation in social networks
  - [Zhang & Yeung, UAI' 2010]
    Multi-Domain Collaborative Filtering
  - [Pan et al, AAAI' 2010]: Transfer Learning in Collaborative Filtering for Sparsity Reduction
- Isolated case
  - [Li et al, IJCAI' 2009]: Can movies and books collaborate? - cross domain collaborative filtering for sparsity reduction
  - [Li et al, ICML' 2009]: Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model

Multi-domain CF [zhang & Yeung *UAI' 2010*]

- For each domain

$$\boxed{\begin{aligned} R^{(i)} &= U^{(i)} V^{(i)\top} + E^{(i)} \\ U_u^{(i)} &\sim \mathcal{N}(0, \sigma^2 I) \\ V_v^{(i)} &\sim \mathcal{N}(0, \sigma^2 I) \\ e^{(i)} &\sim \mathcal{N}(0, \sigma^2) \end{aligned}}$$

- Correlated profile

$$[U^{(i)}] \sim \mathcal{N}(0, \Sigma)$$

$$(i.e.,\ [U_{u\cdot}^{(i)})] \sim \mathcal{N}(0, \Sigma)$$

Multi-domain CF [zhang & Yeung *UAI' 2010*]

- Results

Table 1: Comparison of different methods on the MovieLens data.

Method	1st domain	2nd domain	3rd domain	4th domain	5th domain	Total
PMF	0.9642	1.2104	0.9377	1.0035	1.0352	1.0092
CMF	0.8272	0.7977	0.8120	0.7945	0.7987	0.8088
MCF	0.8061	0.7914	0.7907	0.7761	0.7859	0.7913
MCF-LF	**0.8017**	**0.7644**	**0.7806**	**0.7607**	**0.7504**	**0.7755**

Table 3: Mean of correlation matrix learned by MCF-LF on the MovieLens data in different domains. 1st domain: 'Comedy'; 2nd domain: 'Romance'; 3rd domain: 'Drama'; 4th domain: 'Action'; 5th domain: 'Thriller'.

	1st	2nd	3rd	4th	5th
1st	1.0000	0.8837	0.8584	0.8319	0.8302
2nd	0.8837	1.0000	0.9288	0.8855	0.8805
3rd	0.8584	0.9288	1.0000	0.8647	0.8783
4th	0.8319	0.8855	0.8647	1.0000	0.9122
5th	0.8302	0.8805	0.8783	0.9122	1.0000

Representative work

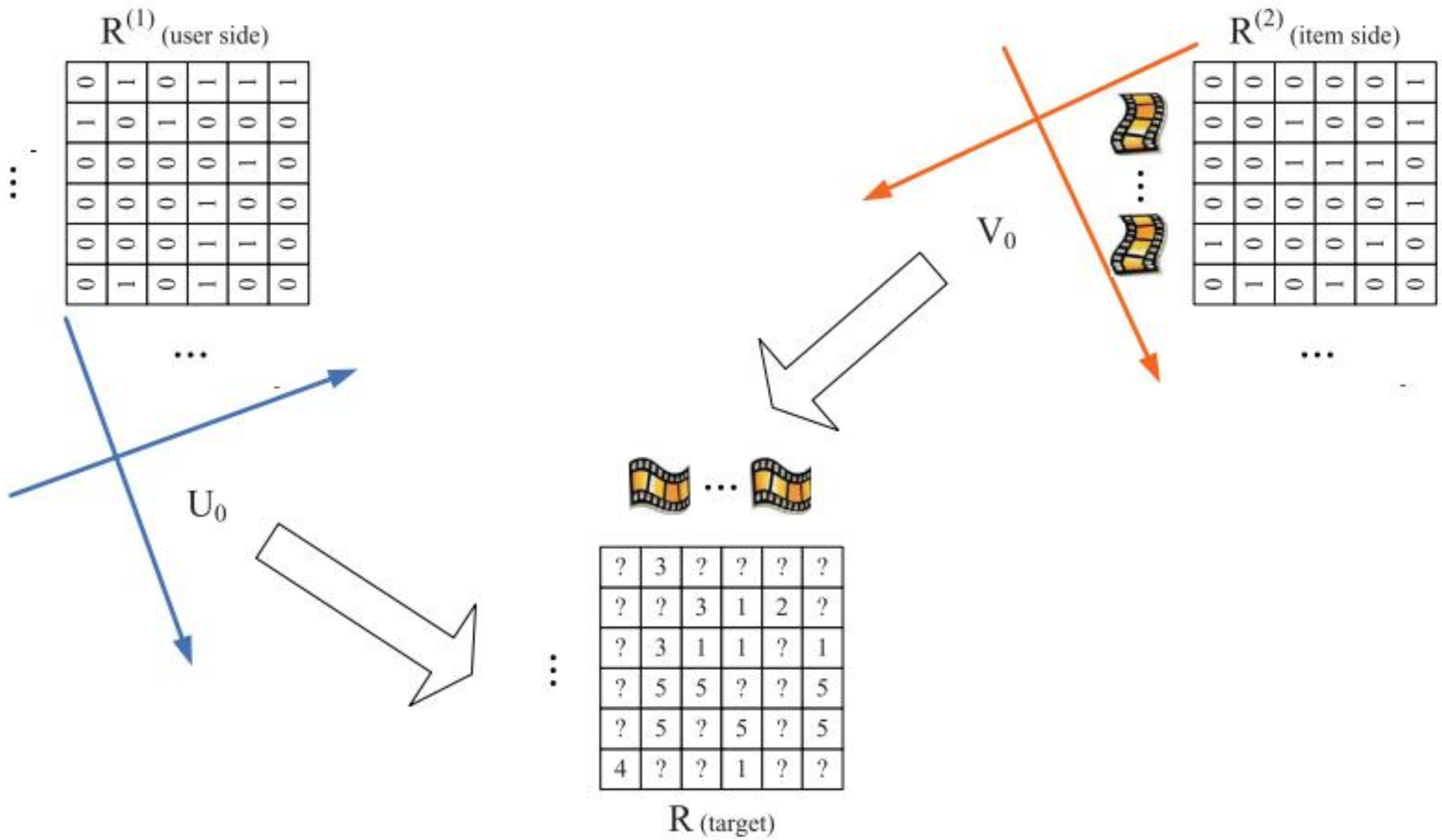
- Shared case
  - [Yang et al, WWW' 2011]

    Like like alike -- Joint friendship and interest propagation in social networks
  - [Zhang & Yeung, UAI' 2010]

    Multi-Domain Collaborative Filtering
  - [Pan et al, AAAI' 2010]: Transfer Learning in Collaborative Filtering for Sparsity Reduction
- Isolated case
  - [Li et al, IJCAI' 2009]: Can movies and books collaborate? - cross domain collaborative filtering for sparsity reduction
  - [Li et al, ICML' 2009]: Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model

Transfer Learning in CF for Sparsity Reduction [Pan et al *AAAI' 2010*]

- Shared users & items



Transfer Learning in CF for Sparsity Reduction [Pan et al *AAAI' 2010*]

- Shared users & items
- Correlated factors ("coordinate system transfer")
  - 1. SVD in R1 and R2 independently

$$\min_{\mathbf{U}^{(i)},\mathbf{V}^{(i)},\mathbf{B}^{(i)}} ||\mathbf{Y}^{(i)} \odot (\mathbf{R}^{(i)} - \mathbf{U}^{(i)}\mathbf{B}^{(i)}\mathbf{V}^{(i)T})||_F^2$$

$$\mathbf{B}^{(i)} = \mathrm{diag}(\sigma_1^{(i)}, \ldots, \sigma_j^{(i)}, \ldots, \sigma_d^{(i)})$$

$$\mathbf{U}^{(i)T}\mathbf{U}^{(i)} = \mathbf{I},\ \mathbf{V}^{(i)T}\mathbf{V}^{(i)} = \mathbf{I}$$

Transfer Learning in CF for Sparsity Reduction [Pan et al *AAAI' 2010*]

- Shared users & items
- Correlated factors ("coordinate system transfer")
  - 1. SVD in R1 and R2 independently
  - 2. Correlated factors (use U1 and V2 as Gaussian means)

$$\min_{\mathbf{U},\mathbf{V},\mathbf{B}} \ ||\mathbf{Y} \odot (\mathbf{R} - \mathbf{U}\mathbf{B}\mathbf{V}^T)|| + \frac{\rho_u}{2}||\mathbf{U} - \mathbf{U}_0||_F^2 + \frac{\rho_v}{2}||\mathbf{V} - \mathbf{V}_0||_F^2$$

$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}, \mathbf{V}^T\mathbf{V} = \mathbf{I}$$

Transfer Learning in CF for Sparsity Reduction [Pan et al *AAAI' 2010*]

- Results

Table 2: Prediction performance of average filling (AF), latent factorization model (LFM), collective matrix factorization (CMF), and coordinate system transfer (CST). Numbers in boldface (i.e. **0.7481**) are the best results among all methods.

	Observed (sparsity)	Without Transfer AF	Without Transfer LFM	With Transfer CMF	With Transfer CST
MAE	10 (0.2%)	0.7764 ± 0.0008	0.8934 ± 0.0005	0.7642 ± 0.0024	**0.7481** ± 0.0014
	20 (0.4%)	0.7430 ± 0.0006	0.8243 ± 0.0019	0.7238 ± 0.0012	**0.7056** ± 0.0008
	30 (0.6%)	0.7311 ± 0.0005	0.7626 ± 0.0008	0.7064 ± 0.0008	**0.6907** ± 0.0006
	40 (0.8%)	0.7248 ± 0.0004	0.7359 ± 0.0008	0.6972 ± 0.0007	**0.6835** ± 0.0008
RMSE	10 (0.2%)	0.9853 ± 0.0011	1.0830 ± 0.0000	0.9749 ± 0.0033	**0.9649** ± 0.0019
	20 (0.4%)	0.9430 ± 0.0006	1.0554 ± 0.0016	0.9261 ± 0.0014	**0.9059** ± 0.0013
	30 (0.6%)	0.9280 ± 0.0005	0.9748 ± 0.0012	0.9058 ± 0.0009	**0.8855** ± 0.0010
	40 (0.8%)	0.9202 ± 0.0003	0.9381 ± 0.0010	0.8955 ± 0.0007	**0.8757** ± 0.0011
Time Complexity		$O(p)$	$O(kpd^2 + k\max(n,m)d^3)$	$O(kpd^2 + k\max(n,m)d^3)$	$O(kpd^3 + kd^6)$

Representative work

- Shared case
  - [Yang et al, WWW' 2011]

    Like like alike -- Joint friendship and interest propagation in social networks
  - [Zhang & Yeung, UAI' 2010]

    Multi-Domain Collaborative Filtering
  - [Pan et al, AAAI' 2010]: Transfer Learning in Collaborative Filtering for Sparsity Reduction
- Isolated case
  - [Li et al, IJCAI' 2009]: Can movies and books collaborate? - cross domain collaborative filtering for sparsity reduction
  - [Li et al, ICML' 2009]: Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model

Can movie and book collaborate ... [Li et al *IJCAI' 2009*]

- Rating pattern ("codebook") transfer

	a	b	c	d	e	f
1	2	3	3	3	2	?
2	3	1	2	2	?	1
3	3	1	?	2	3	1
4	?	2	1	1	3	2
5	2	3	3	3	2	3
6	3	2	1	?	3	2

movie (auxiliary)

	a	e	b	f	c	d
2	3	?	1	1	2	2
3	3	3	1	1	?	2
1	2	2	3	?	3	3
5	2	2	3	3	3	3
4	?	3	2	2	1	1
6	3	3	2	2	1	?

after permutation

	A	B	C
X	3	1	2
Y	2	3	3
Z	3	2	1

cluster-level rating pattern correspondence

	a	b	c	d	e
1	?	1	3	3	1
2	3	3	2	?	3
3	2	2	?	3	?
4	1	1	3	?	?
5	1	?	?	3	1
6	3	?	2	2	3
7	?	2	3	3	2

book (target)

	c	d	a	b	e
1	3	3	?	1	1
4	3	?	1	1	?
5	?	3	1	?	1
2	2	?	3	3	3
6	2	2	3	?	3
3	?	3	2	2	?
7	3	3	?	2	2

after permutation

	A	B
X	3	1
Y	2	3
Z	3	2

Can movie and book collaborate ... [Li et al *IJCAI' 2009*]

- Rating pattern ("codebook") transfer
  - Bi-cluster users/items in auxiliary domain according to ratings
  - Use cluster centers as "codebook" to recover ratings in target domain
  - Predict rating by train a CF model on the recovered ratings

Can movie and book collaborate ... [Li et al *IJCAI' 2009*]

- Rating pattern ("codebook") transfer
  - Bi-cluster users/items in auxiliary domain according to ratings
  - Use cluster centers as "codebook" to recover ratings in target domain
  - Predict rating by train a CF model on the recovered ratings

$$\boxed{\begin{aligned} &\min_{\mathbf{U}\geq 0,\mathbf{V}\geq 0,\mathbf{S}\geq 0} \left\|\mathbf{X}_{aux} - \mathbf{U}\mathbf{S}\mathbf{V}^\top\right\|_F^2 \\ &\text{s.t.} \quad \mathbf{U}^\top\mathbf{U} = \mathbf{I}, \mathbf{V}^\top\mathbf{V} = \mathbf{I}, \end{aligned}}$$

orthogonal nonnegative tri-factorization

Can movie and book collaborate ... [Li et al *IJCAI' 2009*]

- Rating pattern ("codebook") transfer
  - Bi-cluster users/items in auxiliary domain according to ratings
  - Use cluster centers as "codebook" to recover ratings in target domain
  - Predict rating by train a CF model on the recovered ratings

$$\boxed{\mathbf{B} = [\mathbf{U}_{aux}^\top \mathbf{X}_{aux} \mathbf{V}_{aux}] \oslash [\mathbf{U}_{aux}^\top \mathbf{1}\mathbf{1}^\top \mathbf{V}_{aux}]}$$

$$\min_{\substack{\mathbf{U}_{tgt} \in \{0,1\}^{p \times k} \\ \mathbf{V}_{tgt} \in \{0,1\}^{q \times l}}} \left\| [\mathbf{X}_{tgt} - \mathbf{U}_{tgt} \mathbf{B} \mathbf{V}_{tgt}^\top] \circ \mathbf{W} \right\|_F^2 \quad \text{s.t.} \quad \mathbf{U}_{tgt}\mathbf{1} = \mathbf{1}, \mathbf{V}_{tgt}\mathbf{1} = \mathbf{1},$$

Can movie and book collaborate ... [Li et al *IJCAI' 2009*]

- Rating pattern ("codebook") transfer
  - Bi-cluster users/items in auxiliary domain according to ratings
  - Use cluster centers as "codebook" to recover ratings in target domain
  - Predict rating by train a CF model on the recovered ratings

    - $\tilde{\mathbf{X}}_{tgt} = \mathbf{W} \circ \mathbf{X}_{tgt} + [\mathbf{1} - \mathbf{W}] \circ [\mathbf{U}_{tgt}\mathbf{B}\mathbf{V}_{tgt}^{\top}]$

    - Predict rating by memory based CF

Can movie and book collaborate ... [Li et al *IJCAI' 2009*]

- Results
  - Use Eachmovie to help recommendation on Movielens or BookXing

Table 1: MAE on MovieLens (average over 10 splits)

Training Set	Method	Given5	Given10	Given15
ML100	PCC	0.930	0.883	0.873
	CBS	0.874	0.845	0.839
	WLR	0.915	0.875	0.890
	CBT	**0.840**	**0.802**	**0.786**
ML200	PCC	0.905	0.878	0.878
	CBS	0.871	0.833	0.828
	WLR	0.941	0.903	0.883
	CBT	**0.839**	**0.800**	**0.784**
ML300	PCC	0.897	0.882	0.885
	CBS	0.870	0.834	0.819
	WLR	1.018	0.962	0.938
	CBT	**0.840**	**0.801**	**0.785**

Table 2: MAE on Book-Crossing (average over 10 splits)

Training Set	Method	Given5	Given10	Given15
BX100	PCC	0.677	0.710	0.693
	CBS	0.664	0.655	0.641
	WLR	1.170	1.182	1.174
	CBT	**0.614**	**0.611**	**0.593**
BX200	PCC	0.687	0.719	0.695
	CBS	0.661	0.644	0.630
	WLR	0.965	1.024	0.991
	CBT	**0.614**	**0.600**	**0.581**
BX300	PCC	0.688	0.712	0.682
	CBS	0.659	0.655	0.633
	WLR	0.842	0.837	0.829
	CBT	**0.605**	**0.592**	**0.574**

Representative work

- Shared case
  - [Yang et al, WWW' 2011]
    Like like alike -- Joint friendship and interest propagation in social networks
  - [Zhang & Yeung, UAI' 2010]
    Multi-Domain Collaborative Filtering
  - [Pan et al, AAAI' 2010]: Transfer Learning in Collaborative Filtering for Sparsity Reduction
- Isolated case
  - [Li et al, IJCAI' 2009]: Can movies and books collaborate? - cross domain collaborative filtering for sparsity reduction
  - [Li et al, ICML' 2009]: Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model

CF Task I

	a	b	c	d	e	f
1	?	3	?	3	2	3
2	3	1	2	2	?	1
3	3	?	2	?	3	1
4	3	?	1	1	?	2
5	2	3	3	?	2	?
6	3	2	?	1	3	2

Permute rows & cols →

	a	e	b	f	c	d
2	3	?	1	1	2	2
3	3	3	?	1	2	?
1	?	2	3	3	?	3
5	2	2	3	?	3	?
4	3	?	?	2	1	1
6	3	3	2	2	?	1

→

	A	B	C
I	3	1	2
II	2	3	3
III	3	2	1

Cluster-level Rating Matrix

←

	A	B	C	D
I	3	1	2	1
II	2	3	3	1
III	3	2	1	2
IV	1	1	2	3

CF Task II

	a	b	c	d	e
1	3	3	?	?	1
2	2	?	2	1	?
3	?	1	2	1	1
4	?	3	3	3	1
5	2	?	2	1	?
6	?	1	2	1	3
7	1	2	?	2	2

Permute rows & cols →

	b	d	a	c	e
5	?	1	2	2	?
3	1	1	?	2	1
1	3	?	3	?	1
4	3	3	?	3	1
7	2	2	1	?	2
2	?	1	2	2	?
6	1	1	?	2	3

→

	B	C	D
I	1	2	1
II	3	3	1
III	2	1	2
IV	1	2	3

Cluster-level Rating Matrix

←

	A	B	C	D
I	3	1	2	1
II	2	3	3	1
III	3	2	1	2
IV	1	1	2	3

CF Task III

	a	b	c	d	e	f	g
1	2	1	?	3	3	?	1
2	?	?	3	2	1	2	2
3	2	1	2	?	3	3	?
4	1	3	?	1	2	1	3
5	3	2	3	?	?	2	?

Permute rows & cols →

	a	c	d	f	e	b	g
1	2	?	3	?	3	1	1
3	2	2	?	3	3	1	?
2	?	3	2	2	1	?	2
5	3	3	?	2	?	2	?
4	1	?	1	1	2	3	3

→

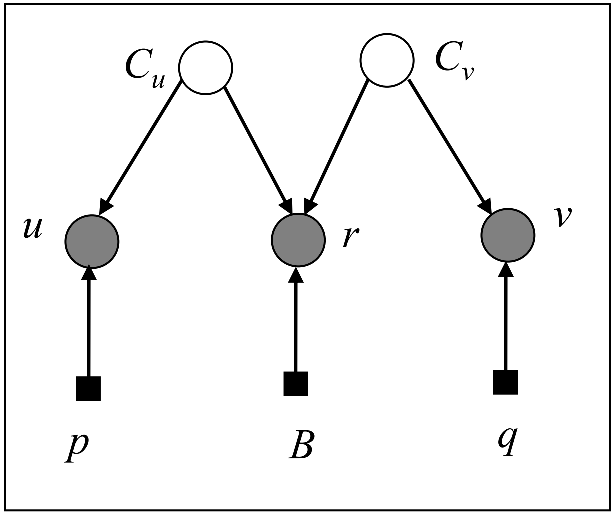
	A	B	C	D
II	2	3	3	1
III	3	2	1	2
IV	1	1	2	3

←

	A	B	C	D
I	3	1	2	1
II	2	3	3	1
III	3	2	1	2
IV	1	1	2	3

Transfer Learning for Collaborative Filtering via... [Li et al *ICML' 2009*]

- Rating pattern ("codebook") transfer
  - A pLSI-style generative model
    - draw a cluster $m \sim p(C_u)$, $n \sim p(C_v)$
    - draw user $u \sim p(u|C_u=m)$, draw item $v \sim p(v|C_v=n)$
    - draw rating $r \sim p(r\,|C_u=m, C_v=n)$



  - Rating prediction

$$f_R(u_i^{(z)}, v_i^{(z)}) = \sum_r r \sum_{k,l} P(r|c_{\mathcal{U}}^{(k)}, c_{\mathcal{V}}^{(l)})\, P(c_{\mathcal{U}}^{(k)}|u_i^{(z)}) P(c_{\mathcal{V}}^{(l)}|v_i^{(z)})$$

$$= \hat{\mathbf{p}}_u^\top \mathbf{B} \hat{\mathbf{p}}_v$$

Transfer Learning for Collaborative Filtering via... [Li et al *ICML' 2009*]

- Results

TRAIN	METHOD	GIVEN5	GIVEN10	GIVEN15
ML100	PCC	0.930	0.908	0.895
	FMM	0.908	0.868	0.846
	RMGM	*0.868*	*0.822*	*0.808*
ML200	PCC	0.934	0.899	0.888
	FMM	0.890	0.863	0.847
	RMGM	*0.859*	*0.821*	*0.806*
ML300	PCC	0.935	0.896	0.888
	FMM	0.885	0.868	0.846
	RMGM	*0.857*	*0.820*	*0.804*
BX100	PCC	0.617	0.599	0.600
	FMM	0.619	0.592	0.583
	RMGM	*0.612*	*0.583*	*0.573*
BX200	PCC	0.621	0.612	0.620
	FMM	0.617	0.602	0.596
	RMGM	*0.615*	*0.591*	*0.583*
BX300	PCC	0.621	0.619	0.630
	FMM	0.615	0.604	0.596
	RMGM	*0.612*	*0.590*	*0.581*

Thanks! 

Any comments would be appreciated!