LITTLE IS MUCH: BRIDGING **CROSS-PLATFORM BEHAVIORS** THROUGH **OVERLAPPED CROWDS**

Meng Jiang (Office 2130SC), Peng Cui (Tsinghua U),
Nicholas Jing Yuan (MSRA), Xing Xie (MSRA),
Shiqiang Yang (Tsinghua U)

cross-platform behavior prediction

AAAI-16 THIRTIETH AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE

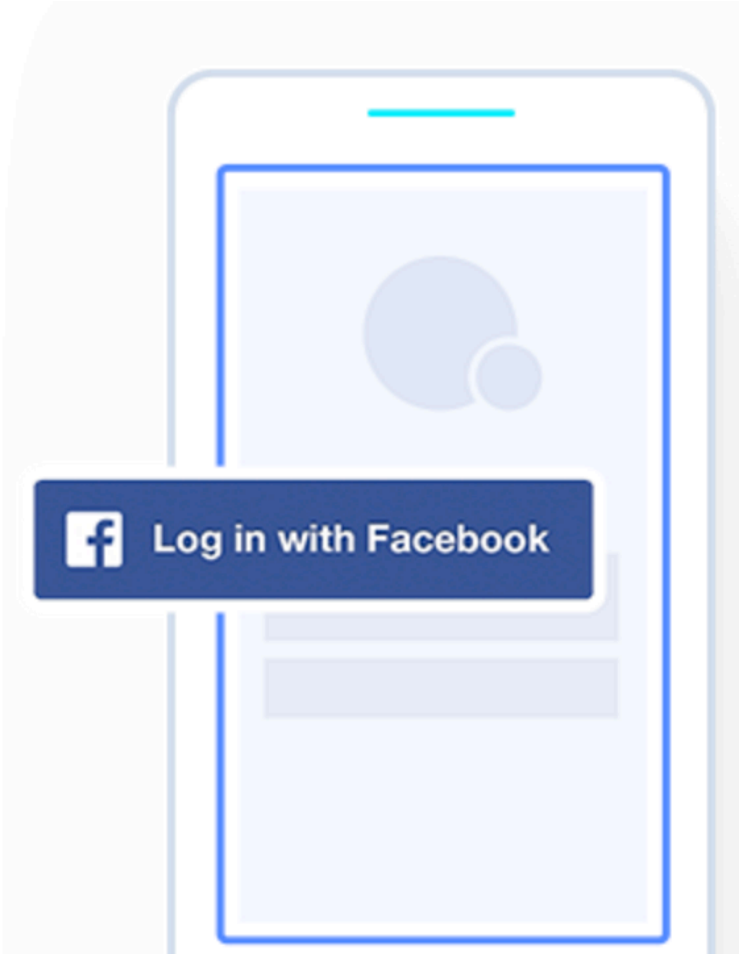
We Need Multiple Platforms

Cross-Platform: e.g., Facebook+App

Add Facebook Login to Your App or Website

Facebook Login for Apps is a secure, fast and convenient way for people to log into your app or website.

- iOS
- Android
- Websites or mobile websites
- More platforms

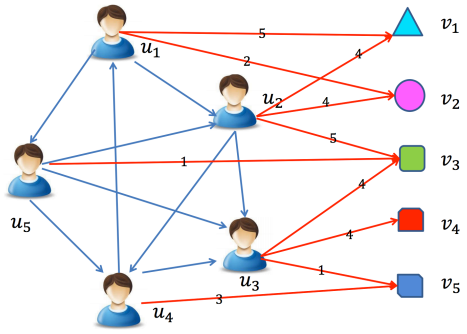


Behavior Prediction

- Predict missing entries in behavioral data
- Applications
  - Uber, Netflix, Amazon…
  - Recommender systems
  - Target advertising

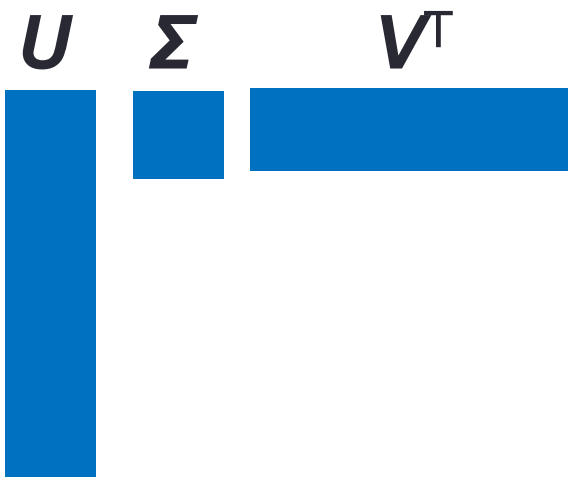
Matrix Factorization

- Low-rank MF on the user-item rating matrix ***R***
- User preference vector ***U*** and Item vector ***V***
- High sparsity is the key challenge
- If we have auxiliary data…



R

	v_1	v_2	v_3	v_4	v_5
u_1	5	?	2	?	?
u_2	4	4	5	?	?
u_3	?	?	4	4	1
u_4	?	?	?	?	3
u_5	?	?	1	?	?



Koren. **Factorizatoin Meets the Neighborhood: A Multifaceted Collaborative Filtering Model.** *KDD,* **2008.**

Knowledge Transfer: Alleviate the Sparsity

Problem	Bridge	Method
Cross-Source		
Cross-Domain		
Cross-Platform		

Cross Data Sources

- Non-overlapped: Codebook Transfer (CBT)

$k_{user} \times k_{item}$

Li et al. **Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.** *IJCAI*, 2009.

Cross Data Sources

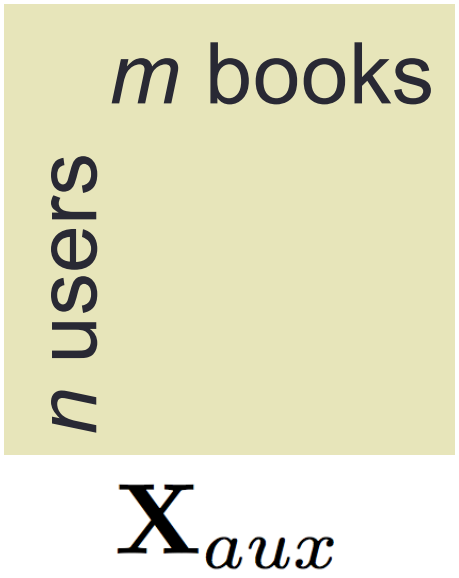
- Non-overlapped: Codebook Transfer (CBT)

$k_{user} \times k_{item}$

Li et al. **Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.** *IJCAI*, 2009.

Cross Data Sources

- Codebook Transfer (CBT)



$n \times k$ $m \times l$

$$\min_{\mathbf{U}\geq 0,\mathbf{V}\geq 0,\mathbf{S}\geq 0} \left\|\mathbf{X}_{aux} - \mathbf{U}\mathbf{S}\mathbf{V}^\top\right\|_F^2$$

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

Codebook = User Group × Item Group

$k \times l$

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

Li et al. **Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.** *IJCAI*, 2009.

Cross Data Sources

- Codebook Transfer (CBT)

$\mathbf{X}_{tgt}$

Codebook: k × l

p × k

q × l

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

$$\text{s.t.} \quad \mathbf{U}_{tgt}\mathbf{1}=\mathbf{1}, \mathbf{V}_{tgt}\mathbf{1}=\mathbf{1},$$

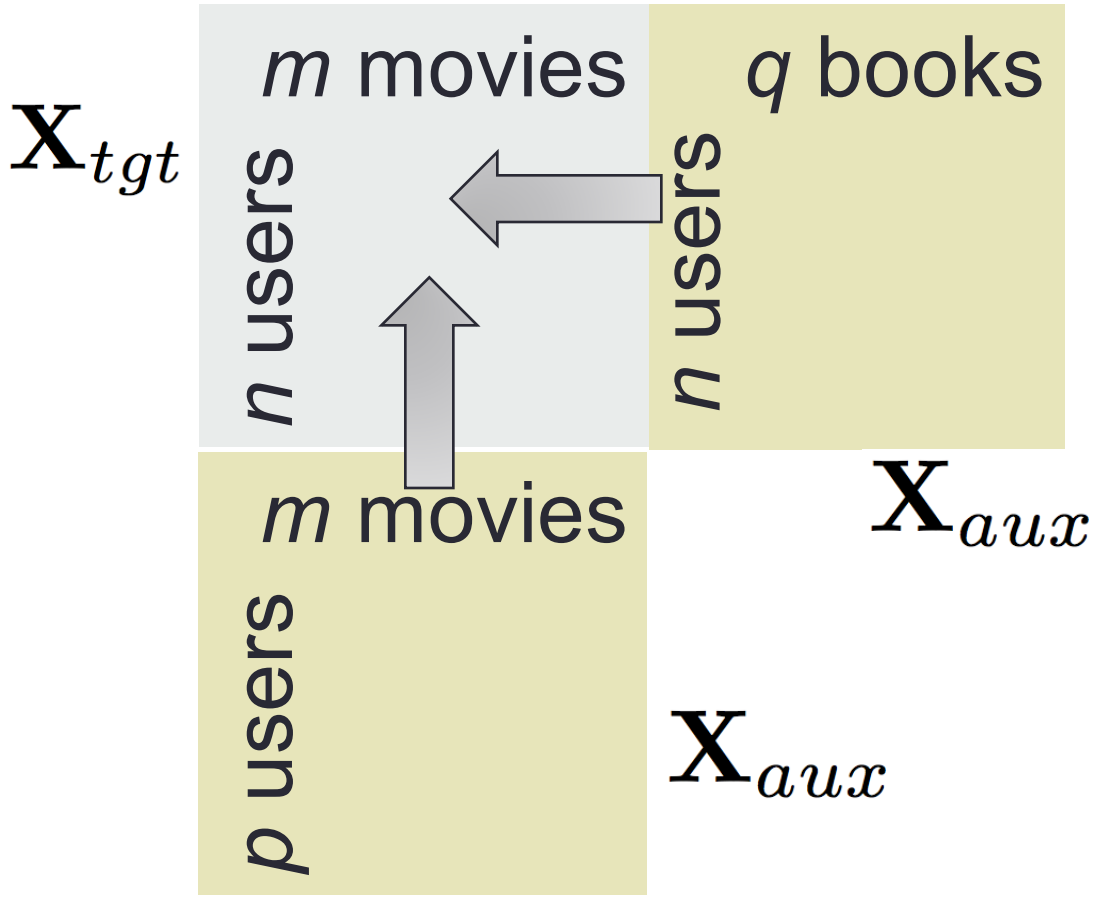
Li et al. **Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.** *IJCAI*, 2009.

Knowledge Transfer: Alleviate the Sparsity

Problem	Bridge	Method
Cross-Source	Non-overlapped → User cluster × Item cluster → The same latent representation	CBT
Cross-Domain		
Cross-Platform		

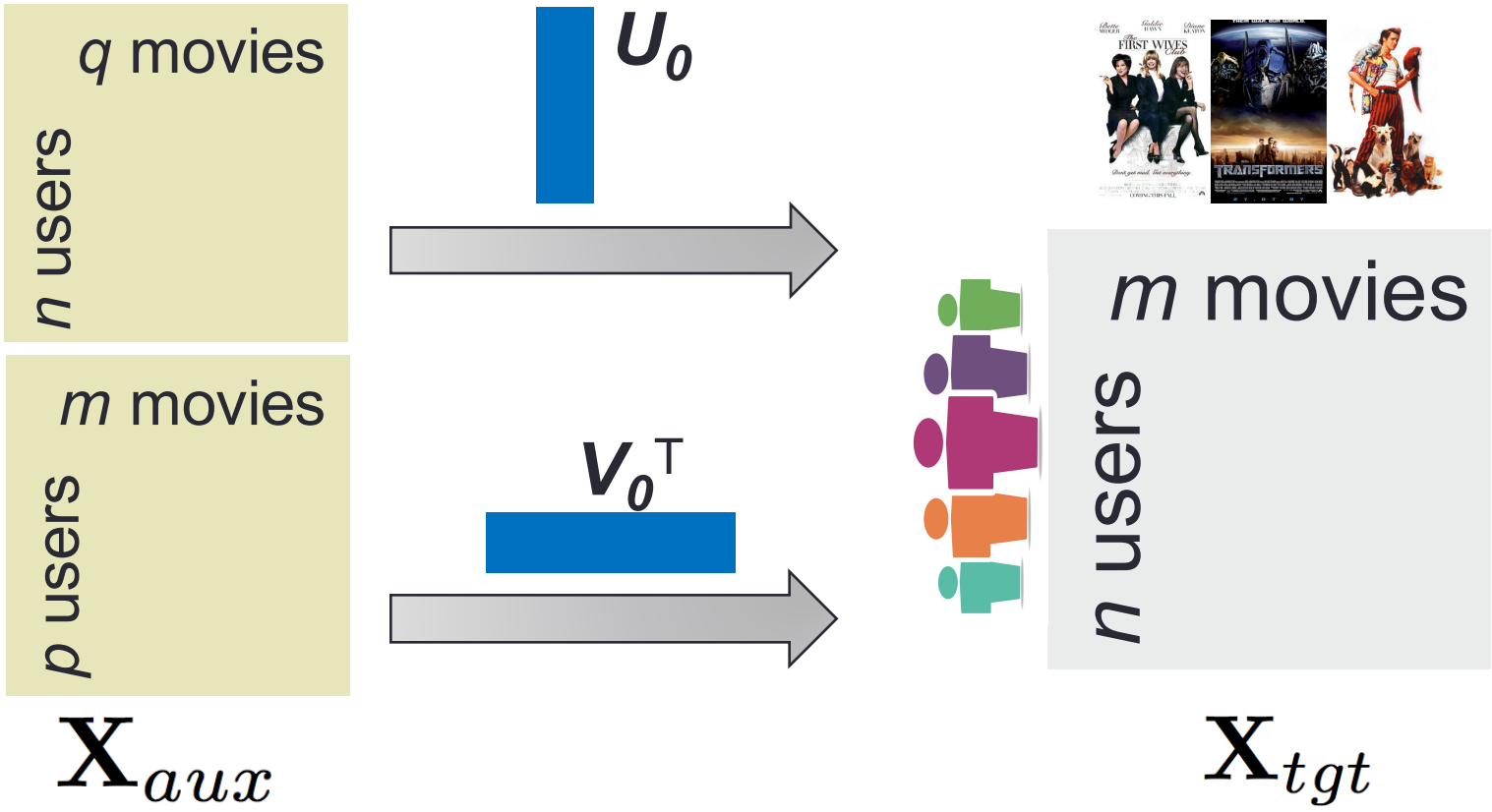
Cross Domains

- Fully overlapped



Cross Domains

- Coordinate System Transfer (CST)



Pan et al. **Transfer Learning in Collaborative Filtering for Sparsity Reduction.** *AAAI*, 2010.

Cross Domains

- Coordinate System Transfer (CST)

Auxiliary data:

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

Target data:

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

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

Pan et al. **Transfer Learning in Collaborative Filtering for Sparsity Reduction.** *AAAI*, 2010.

Knowledge Transfer: Alleviate the Sparsity

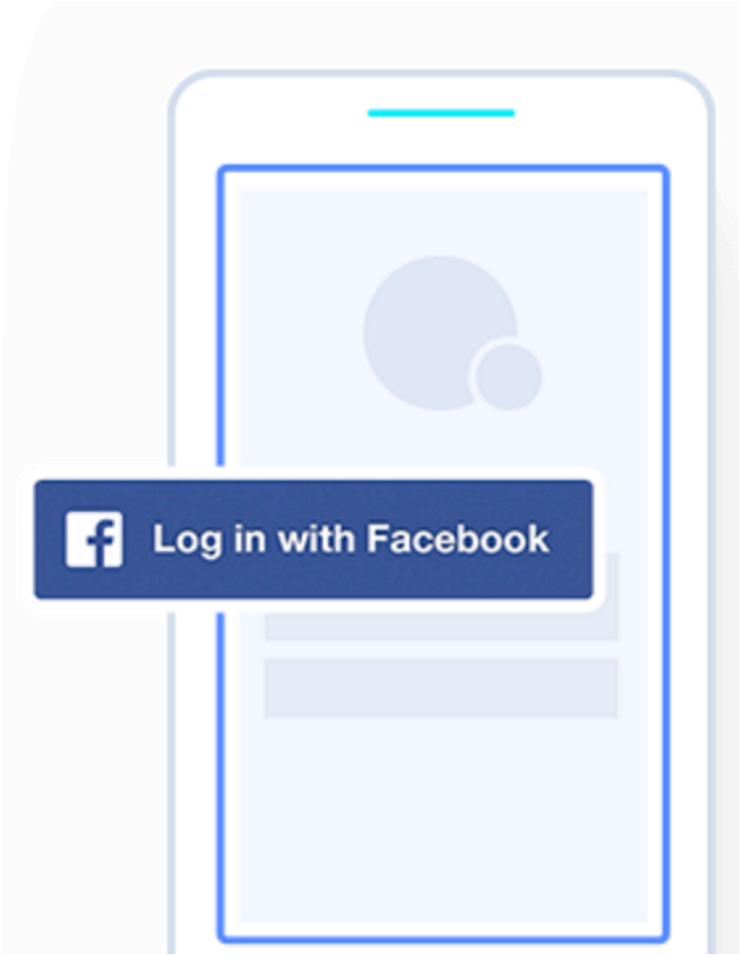
Problem	Bridge	Method
Cross-Source	Non-overlapped → User cluster × Item cluster → The same latent representation	CBT
Cross-Domain	Fully overlapped users OR Fully overlapped items → User vector OR Item vector → The same latent representation	CST
Cross-Platform		

Cross-Platform: e.g., Facebook+App

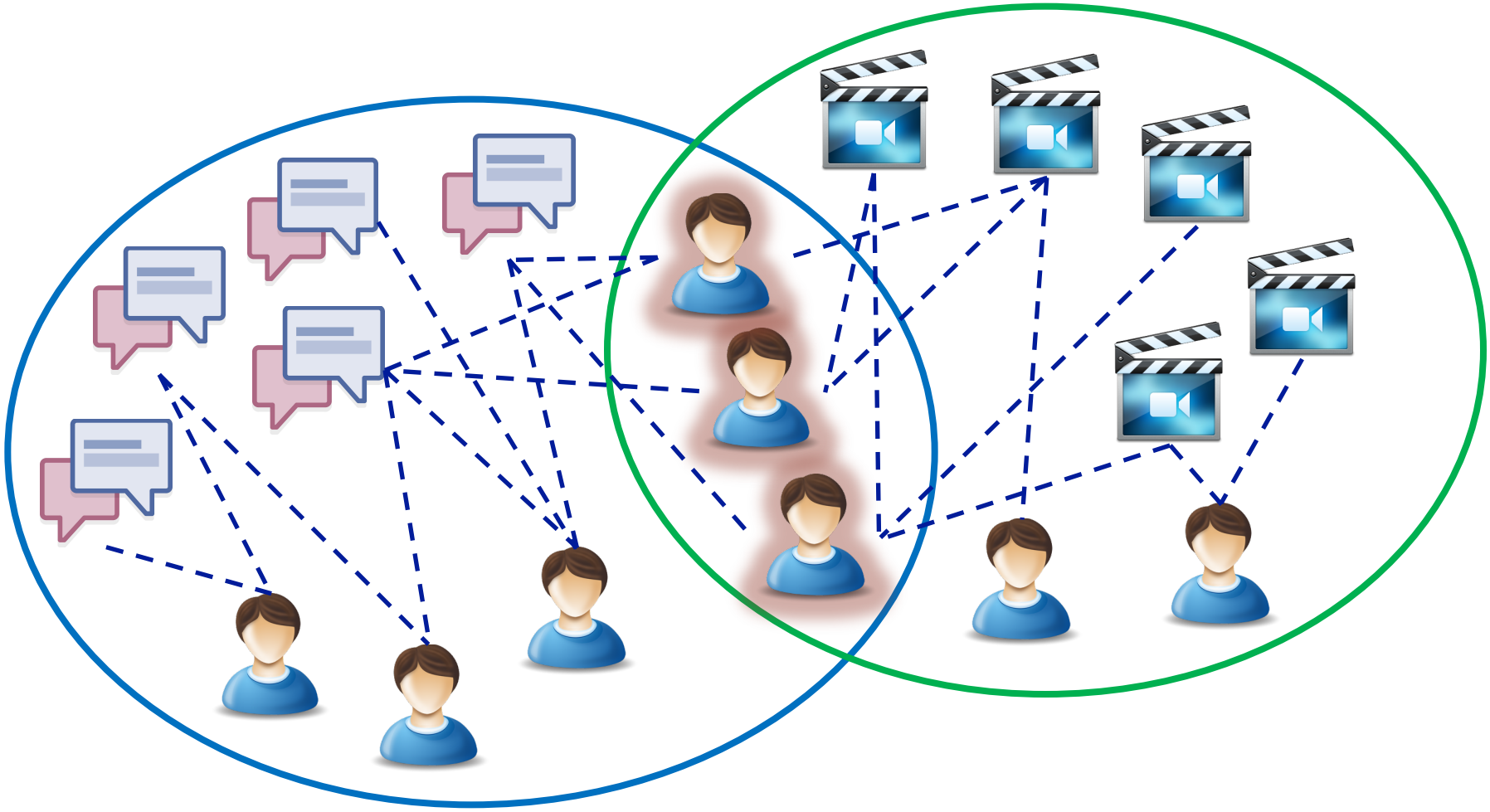
Add Facebook Login to Your App or Website

Facebook Login for Apps is a secure, fast and convenient way for people to log into your app or website.

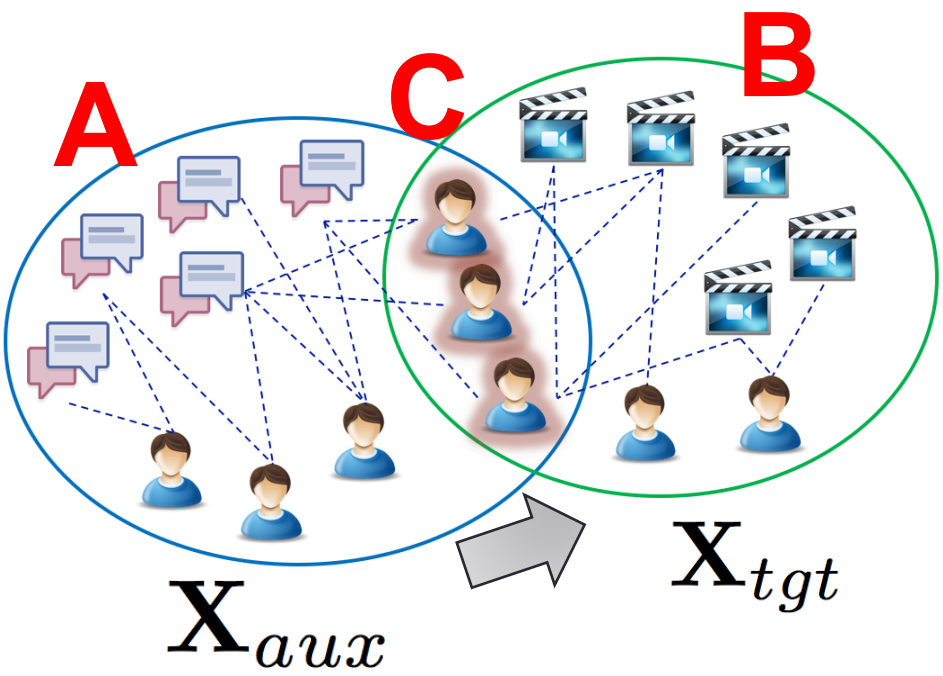
- iOS
- Android
- Websites or mobile websites
- More platforms

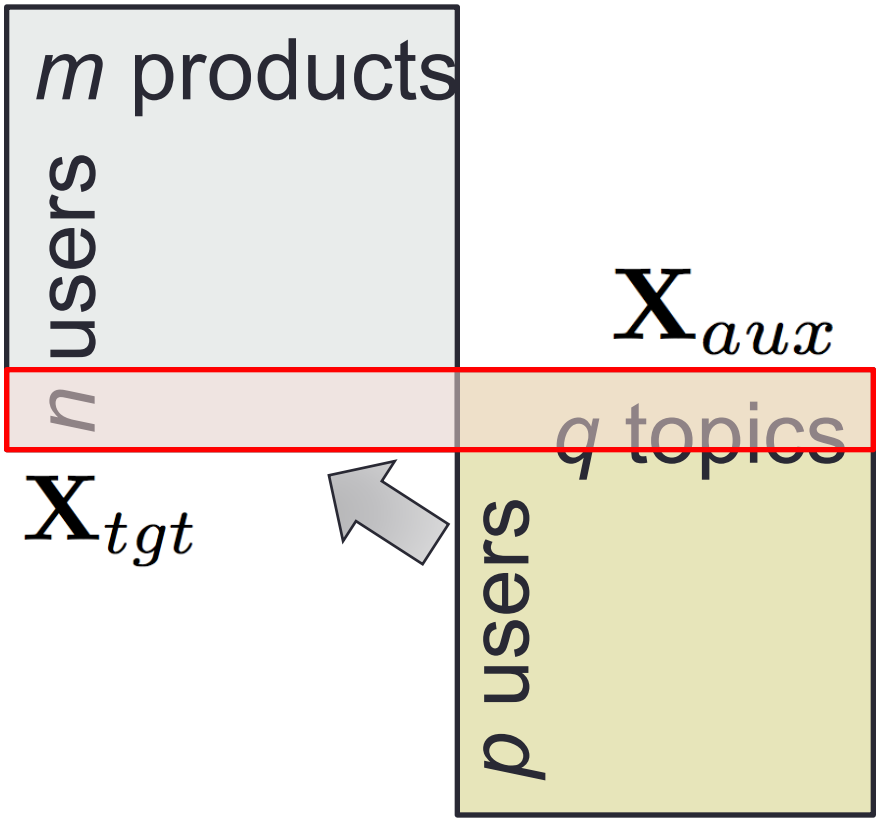


Facebook/Twitter/Google Chrome (Auxiliary)
Uber/Amazon/Netflix (Target)



Partially Overlapped Users





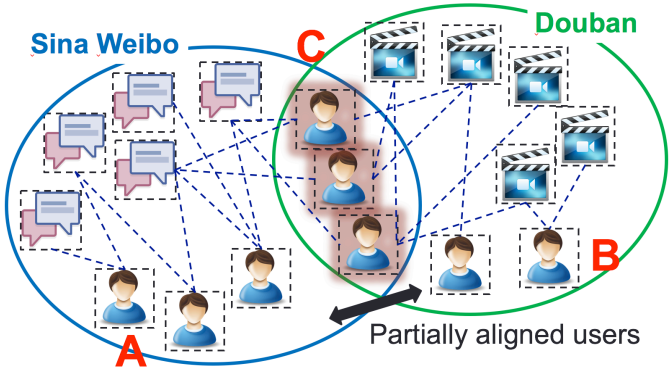
Knowledge Transfer: Alleviate the Sparsity

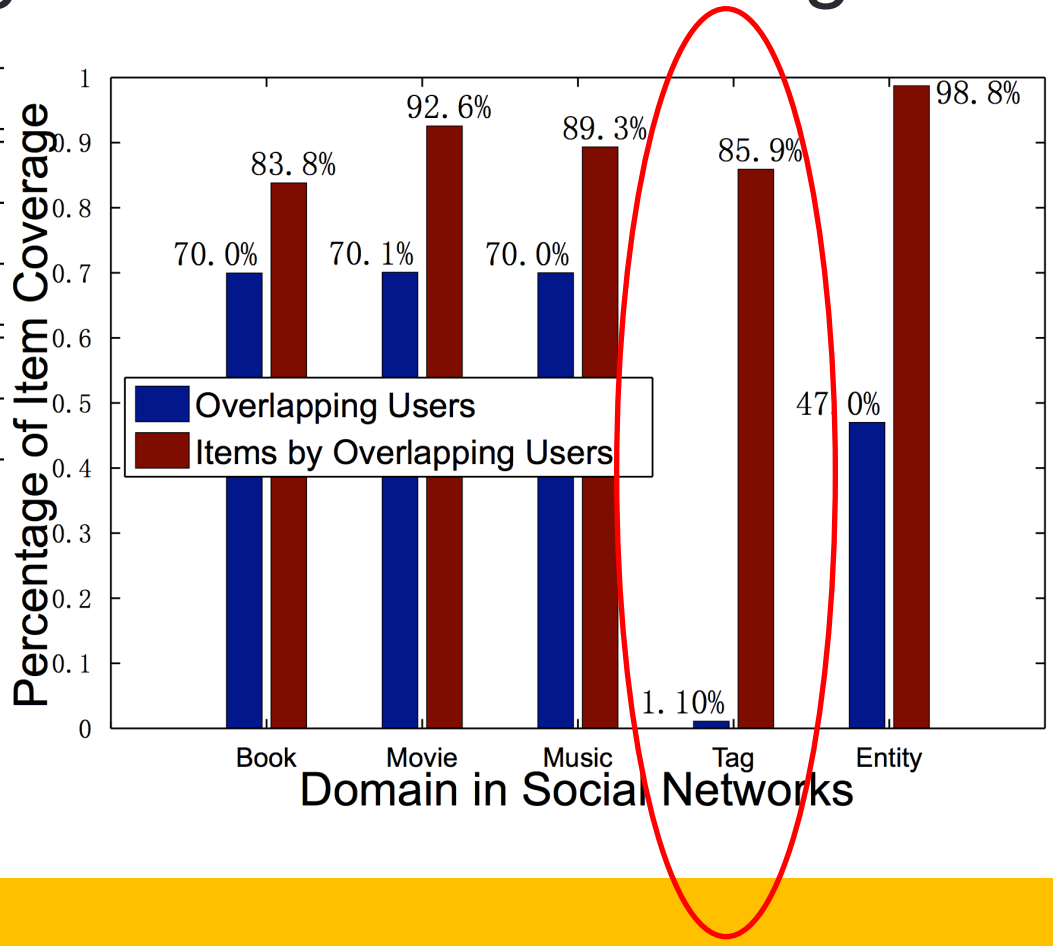
Problem	Bridge	Method
Cross-Source	Non-overlapped → User cluster × Item cluster → The same latent representation	CBT
Cross-Domain	Fully overlapped users OR Fully overlapped items → User vector OR Item vector → The same latent representation	CST
Cross-Platform	**Partially overlapped users** **→ ?**	**?**

If Overlapped Users Adopted Many Items

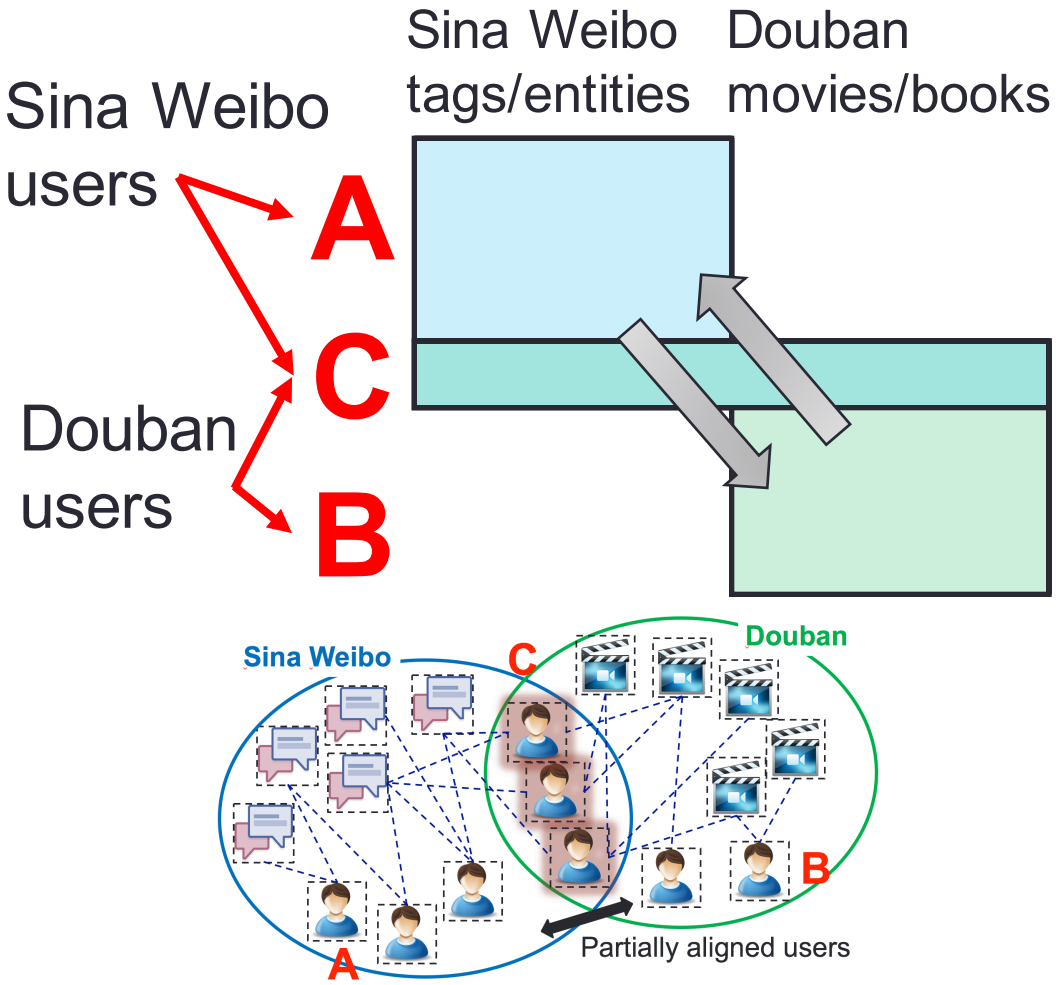
- 1.1% Overlapping Crowd: 85% Coverage

	#User	#Item
Book	30,536	212,835
Movie	40,246	64,090
Music	33,938	286,464
Social tag	2,721,365	10,176
Tweet entity	25,586	113,591





When NO Transfer



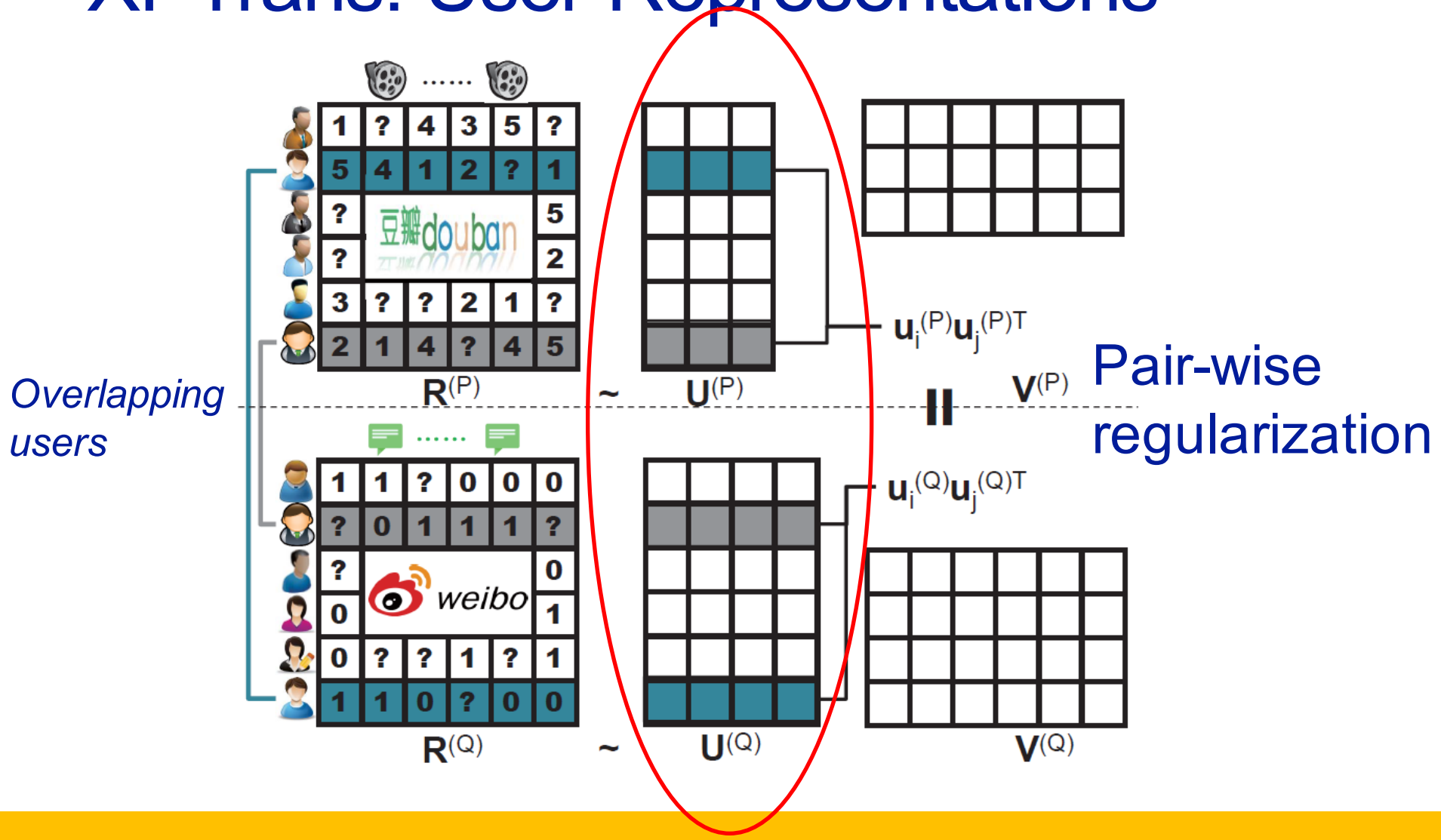
User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A	Auxiliary platform data!	
C	**0.779**	**0.805**
B	1.439	0.640

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	0.429	0.464
C	**0.267**	**0.666**
B	Auxiliary platform data!	

Knowledge Transfer: Alleviate the Sparsity

Problem	Bridge	Method
Cross-Source	Non-overlapped → User cluster × Item cluster → The same latent representation	Codebook
Cross-Domain	Fully overlapped users OR Fully overlapped items → User vector OR Item vector → The same latent representation	CST
Cross-Platform	**Partially overlapped users** **→ User vector** **→ Different latent representations**	**XPTrans**

XPTrans: User Representations



XPTrans: Semi-Supervised Transfer

- Input
  - Tgt./Aux. platform P/Q;
  - Behavior data $R^{(P)}$/$R^{(Q)}$;
  - Observation $W^{(P)}$/$W^{(Q)}$;
  - Overlapping indicator $W^{(P,Q)}$,
- Output
  - User latent representation $U^{(P)}$/$U^{(Q)}$;
  - Item latent representation $V^{(P)}$/$V^{(Q)}$;
  - Missing values in $R^{(P)}$

- Objective function

Target platform

Auxiliary platform

$$\mathcal{J} = \sum_{i,j} W_{i,j}^{(P)} \left(R_{i,j}^{(P)} - \sum_{r} U_{i,r}^{(P)} V_{r,j}^{(P)} \right)^2 + \lambda \sum_{i,j} W_{i,j}^{(Q)} \left(R_{i,j}^{(Q)} - \sum_{r} U_{i,r}^{(Q)} V_{r,j}^{(Q)} \right)^2 + \mu \sum_{i_1,j_1,i_2,j_2} W_{i_1,j_1}^{(P,Q)} W_{i_2,j_2}^{(P,Q)} \left(A_{i_1,i_2}^{(P)} - A_{j_1,j_2}^{(Q)} \right)^2$$

Overlapping user similarity (Pair-wise regularization)

Leveraging Auxiliary Platform Data

NO Transfer

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A	Auxiliary platform data!	
C	0.779	0.805
B	**1.439**	**0.640**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.429**	**0.464**
C	0.267	0.666
B	Auxiliary platform data!	

Transfer via the Same Latent Space

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A		
C	0.757	0.811
B	**1.164 (-19%)**	**0.702 (+9.7%)**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.411 (-4.2%)**	**0.487 (+5.0%)**
C	0.256	0.681
B		

Leveraging Flexible Representation

Transfer via the Same Latent Space

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A		
C	0.757	0.811
B	**1.164**	**0.702**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.411**	**0.487**
C	0.256	0.681
B		

Transfer via Different Latent Spaces

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A		
C	0.715	0.821
B	**0.722 (-38%)**	**0.820 (+17%)**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.374 (-11%)**	**0.533 (+12%)**
C	0.236	0.705
B		

Cross-Platform: Where Amazing Happens

NO Transfer

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A	Auxiliary platform data!	
C	0.779	0.805
B	**1.439**	**0.640**

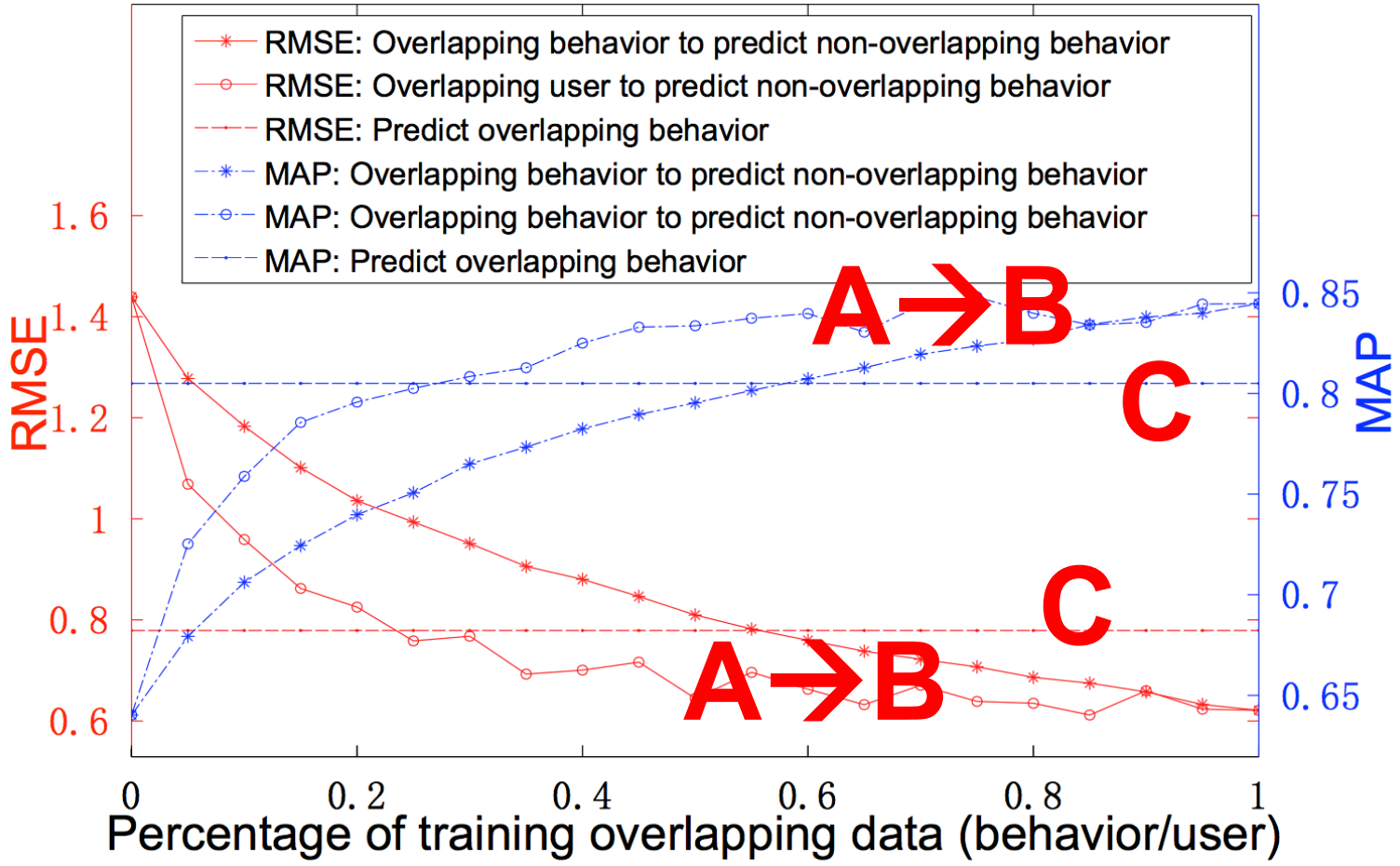
User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.429**	**0.464**
C	0.267	0.666
B	Auxiliary platform data!	

Transfer via Different Latent Spaces

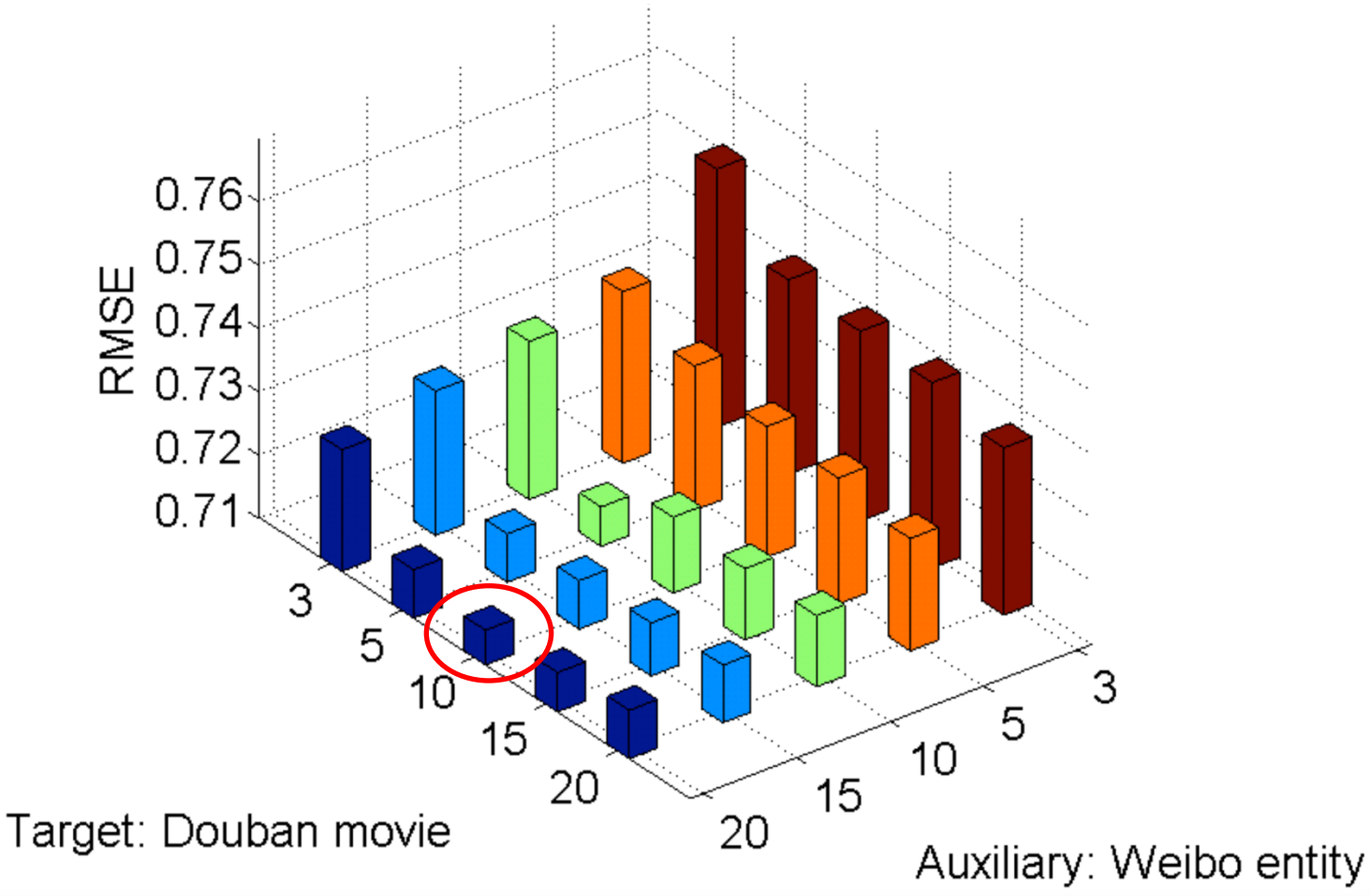
User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A		
C	0.715	0.821
B	**0.722**	**0.820**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.374**	**0.533**
C	0.236	0.705
B		

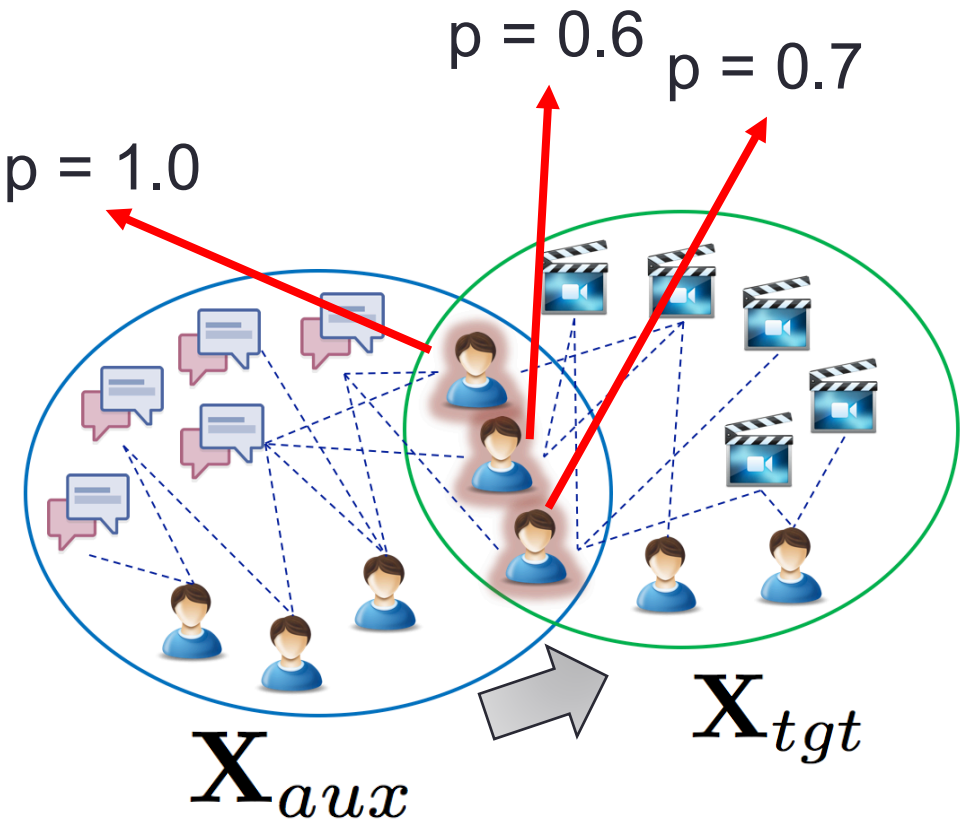
Performance



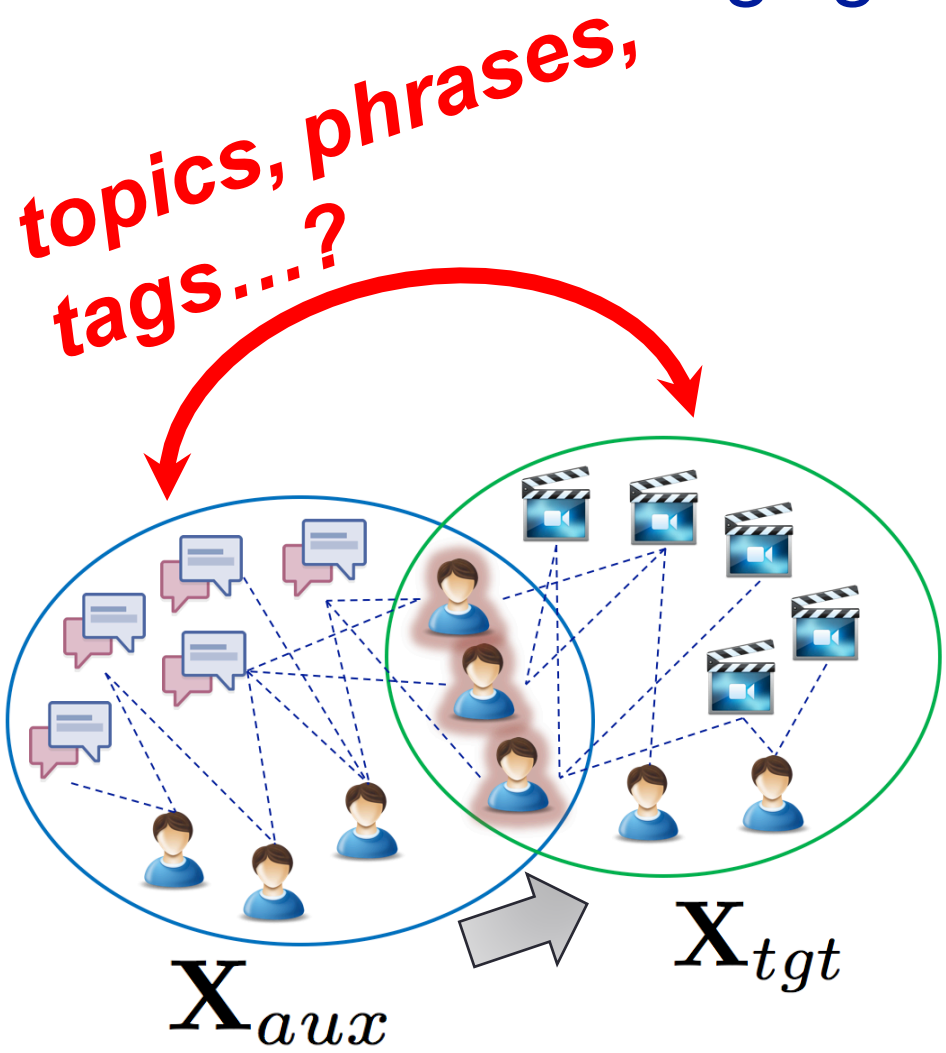
Different Sizes of Latent Space

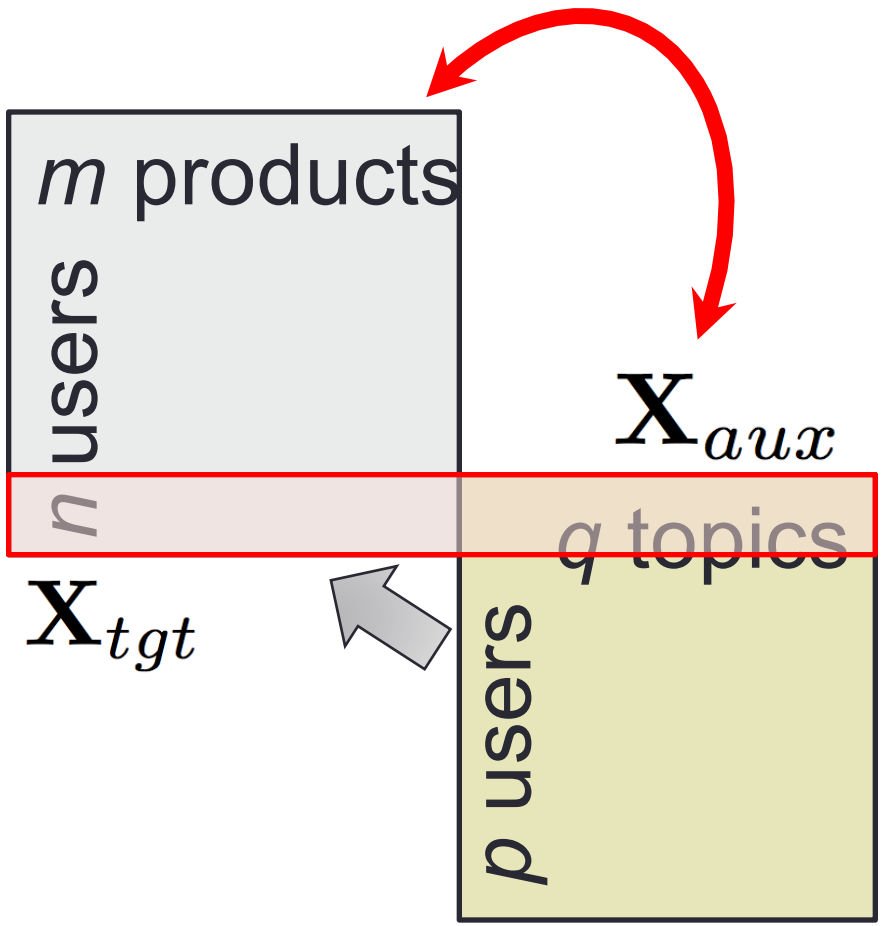


Discussion 1: Probability of the Overlapping?



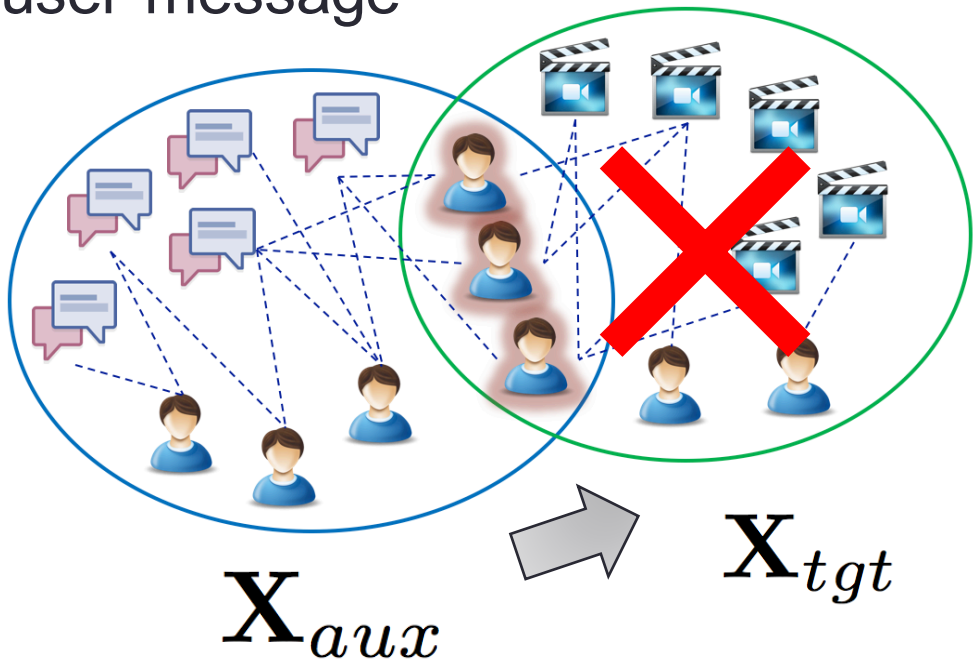
Discussion 2: Bridging Items with Semantics?





Discussion 3:
Bridging Different Behavioral Data Structures?

Facebook
user-message



Uber:
spatio-temporal data
Pinterest:
multimedia data
Amazon:
user-product perchasing
Netflix:
user-movie rating
….

THANK YOU

www.meng-jiang.com