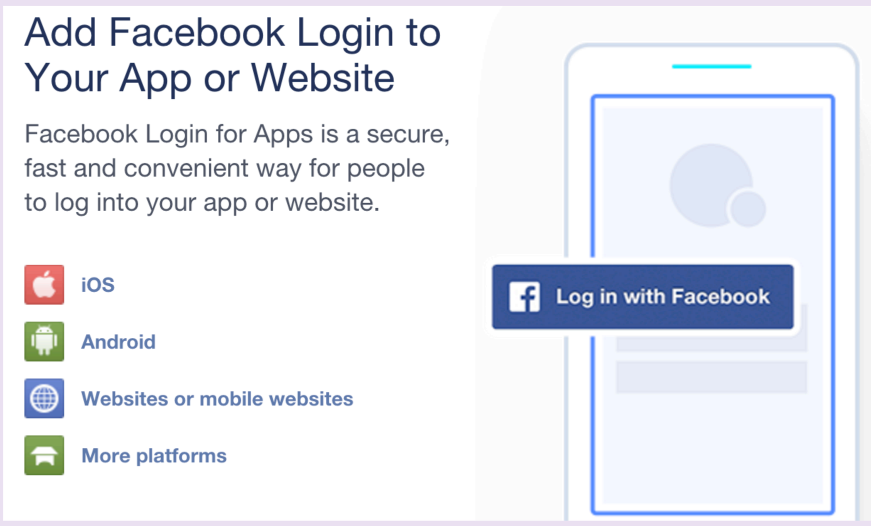
Little is Much: Bridging Cross-Platform Behaviors through Overlapped Crowds

XPTrans

Meng Jiang[1], Peng Cui[1], Nicholas Jing Yuan[2], Xing Xie[2], Shiqiang Yang[1]
[1] Department of Computer Science and Technology, Tsinghua University, Beijing, China
[2] Microsoft Research Asia, Beijing, China

Cross-Platform Behavior Prediction

We register the Uber application with our Facebook accounts. So, can we improve the behavior prediction accuracy on Uber with the rich social data?

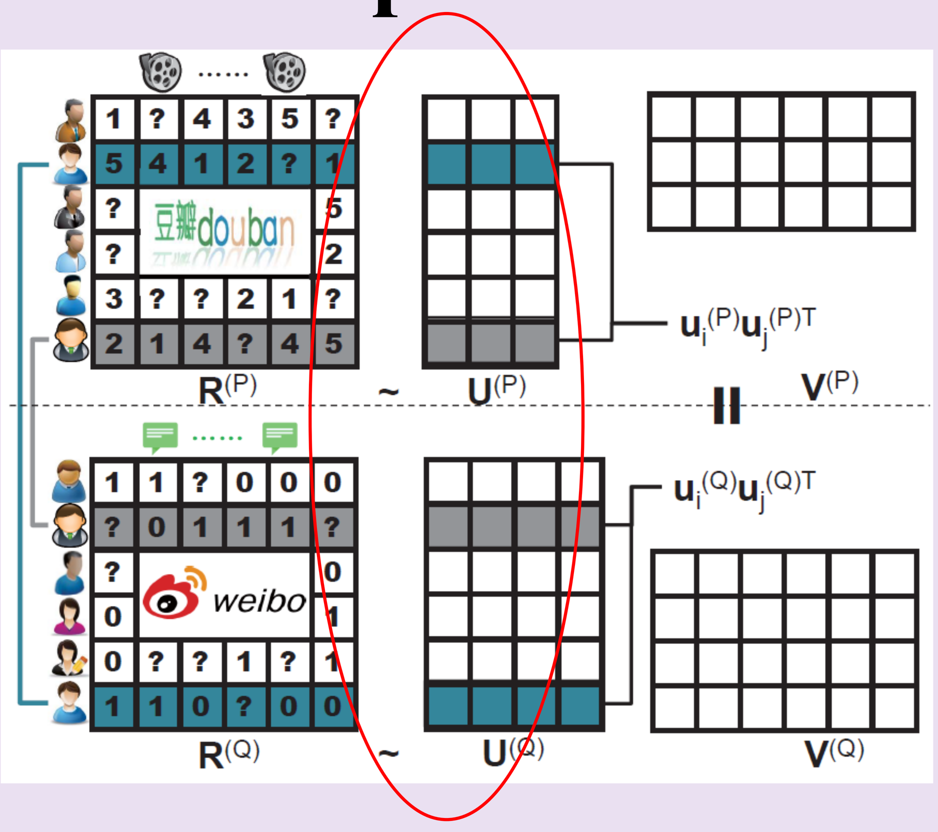


The key challenge in prediction is sparsity. **Knowledge transfer from auxiliary data can alleviate the sparsity problem.**

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 Framework

Semi-supervised Transfer Learning



Overlapping user similarity across platforms as flexible regularization

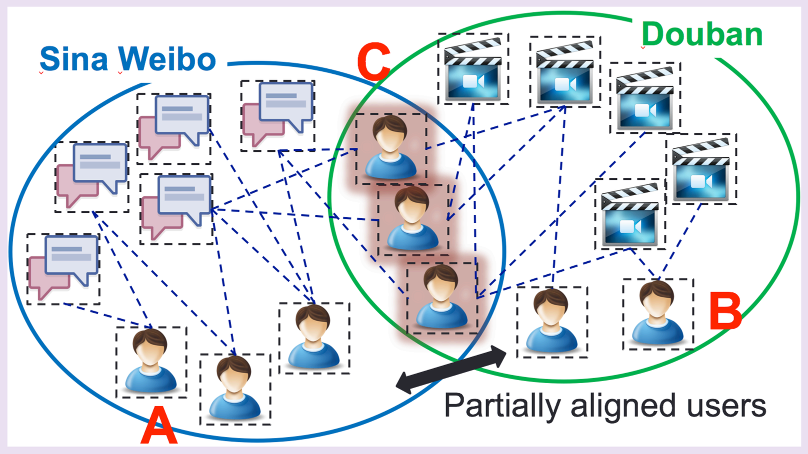
Objective Function

Target platform Auxiliary platform

$$\mathcal{J} = \sum_{i,j} W^{(P)}_{i,j}\left(R^{(P)}_{i,j} - \sum_{r} U^{(P)}_{i,r}V^{(P)}_{r,j}\right)^2 + \lambda \sum_{i,j} W^{(Q)}_{i,j}\left(R^{(Q)}_{i,j} - \sum_{r} U^{(Q)}_{i,r}V^{(Q)}_{r,j}\right)^2 + \mu \sum_{i_1,j_1,i_2,j_2} W^{(P,Q)}_{i_1,j_1}W^{(P,Q)}_{i_2,j_2}\left(A^{(P)}_{i_1,i_2} - A^{(Q)}_{j_1,j_2}\right)^2$$

Overlapping user similarity (Pair-wise regularization)

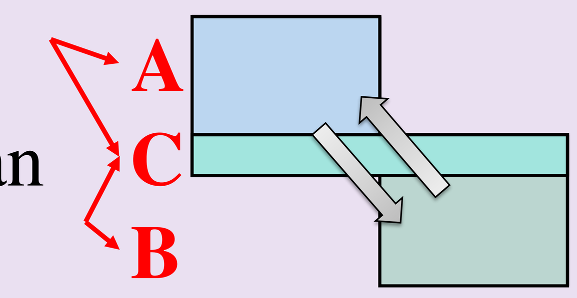
Experiments



Sina Weibo tags/entities Douban movies/books

Sina Weibo users

Douban users



C: Overlapping users; A and B are NOT.

Q: Can we transfer the auxiliary big data A to improve the performance on sparse data B as good as richer but small data C?

Transfer but the Same Latent Space Size

NO Transfer

Weibo tweet entity to Douban movie	RMSE	MAP
A		
C	0.779	0.805
B	1.439	0.640

Douban book to Weibo social tag	RMSE	MAP
A	0.429	0.464
C	0.267	0.666
B		

Transfer via the Same Latent Space

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

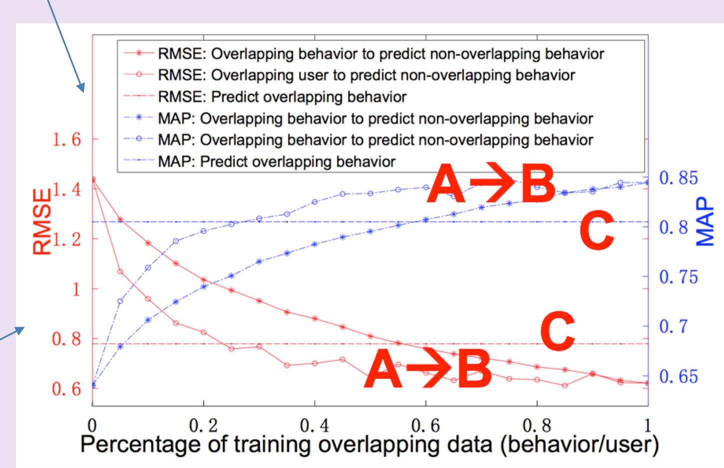
Douban book to Weibo social tag	RMSE	MAP
A	0.411 (-4.2%)	0.487 (+5.0%)
C	0.256	0.681
B		

Transfer via Different Latent Spaces

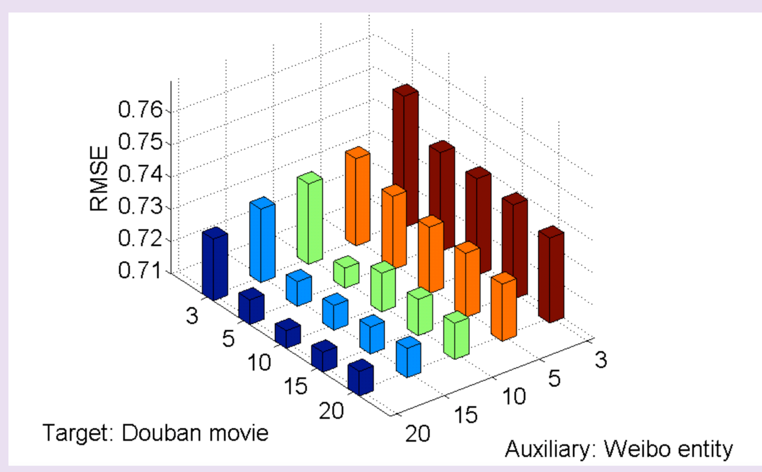
Weibo tweet entity to Douban movie	RMSE	MAP
A		
C	0.715	0.821
B	**0.722 (-38%)**	**0.820 (+17%)**

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

Performance



Different sizes of Latent Space



Acknowledgement

Contact: Meng Jiang
mjiang89@gmail.com
Dr. Jiang is now a postdoc at UIUC.

清華大學 Tsinghua University

Microsoft Research 微软亚洲研究院