STAR: croSs-plaTform App Recommendation

Da Cao, Xiangnan He, Liqiang Nie, Xiaochi Wei, Xia Hu, Shunxiang Wu, Tat-Seng Chua

This work was finished when Da Cao was a Ph.D. student in XMU and a visiting student in NUS.

Currently Da Cao is an assistant professor in HNU.

caoda0721@gmail.com

Submit: June 2016

Accepted: Nov 2016

Published: July 2017

Outline

- Background
- Proposed Method
- Experiments and Results
- Conclusion

App Development

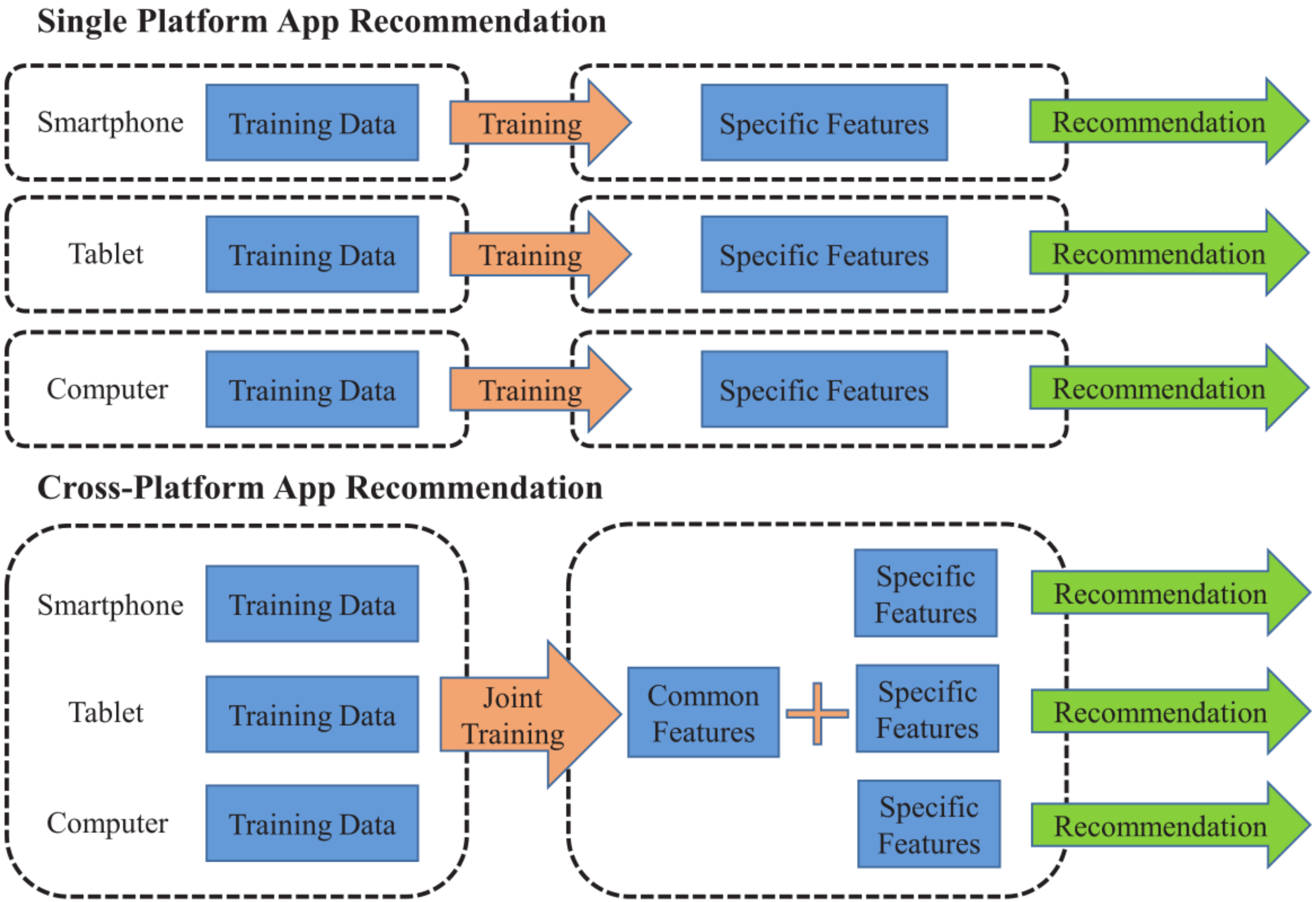
Mobile Network

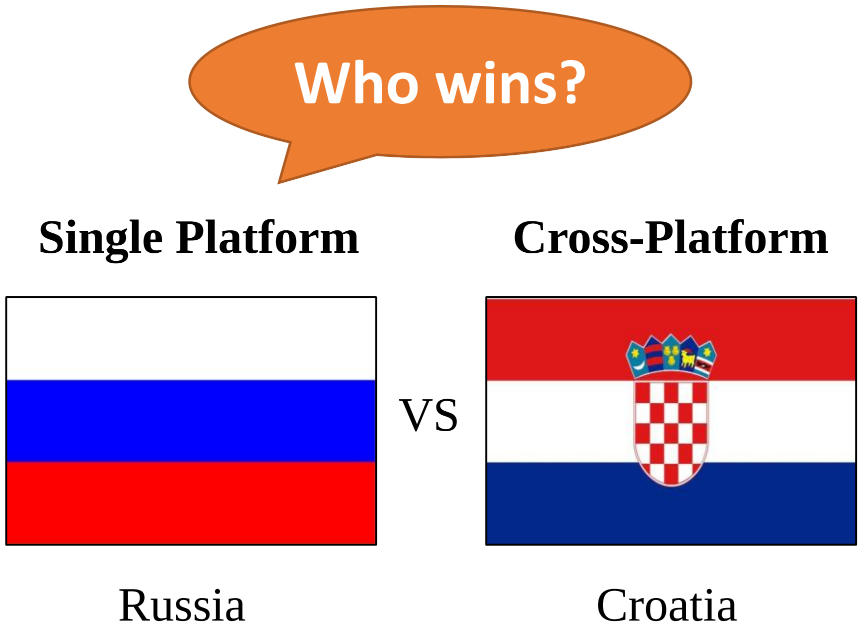
App-Driven Life

Multi-Platform

Overwhelmed

Two App Recommendation solutions





Challenges

- Platform Variance
- Data Heterogeneity
- Data Sparsity
- Cold-Start Problem

Outline

- Background
- Proposed Method
- Experiments and Results
- Conclusion

Proposed Method

$$\hat{r}_{ijs} = \mu + \mathbf{b}_u(i) + \mathbf{b}_v(js) + \mathbf{u}_i^T \mathbf{v}_{js}$$

$$\mathbf{v}_{js} = \mathbf{w}_j + \mathbf{M}\boldsymbol{\theta}_{js}$$

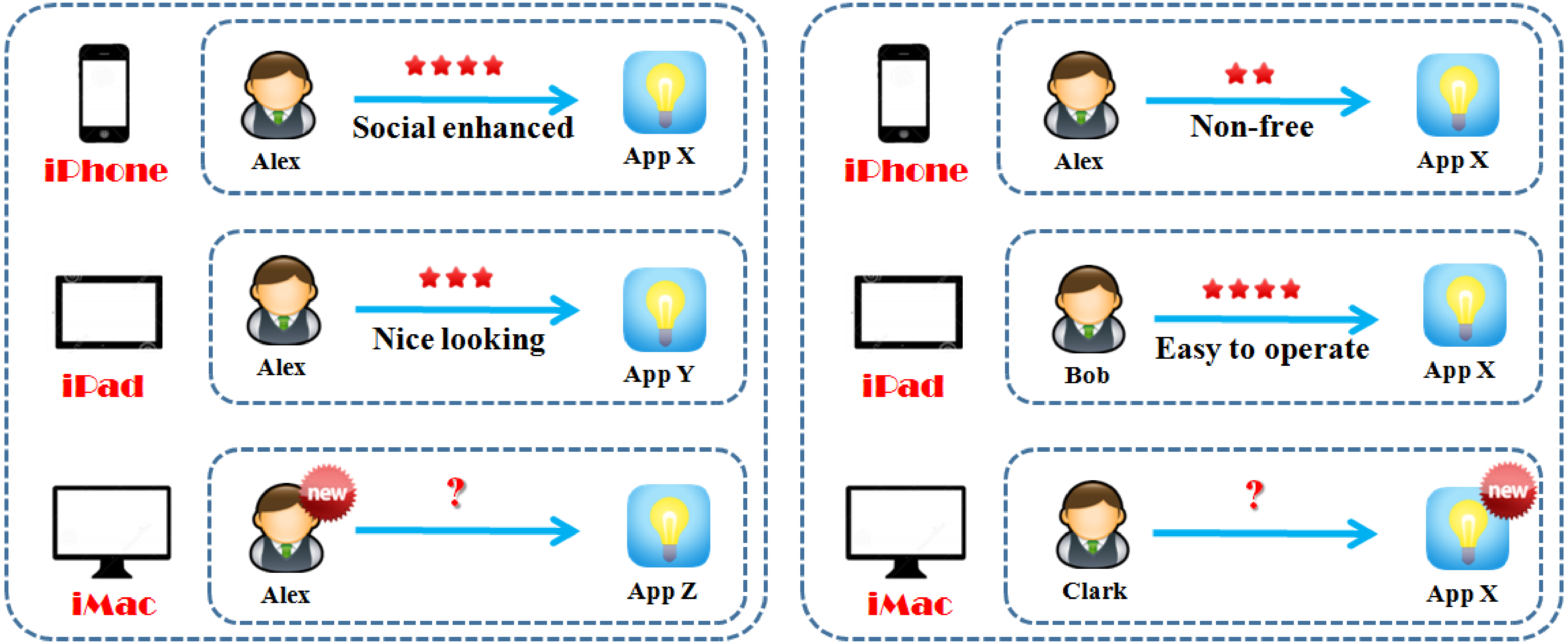
$$\hat{r}_{ijs} = \mu + \mathbf{b}_u(i) + \mathbf{b}_w(j) + \mathbf{u}_i^T(\mathbf{w}_j + \mathbf{M}\boldsymbol{\theta}_{js})$$

$$\mathcal{L}(\Theta) = \frac{1}{2} \sum_{(i,j,s)\in\mathcal{R}} (r_{ijs} - \hat{r}_{ijs})^2 + \lambda||\Theta||^2$$

$$\text{where } \Theta = \{\mathbf{U}, \mathbf{W}, \mathbf{M}, \mathbf{b}_u, \mathbf{b}_w\}$$

Cold-Start Problems

New-user cold-start | **New-App cold-start**



The question mark "?" stands for the ratings that we wish to predict, and the label "new" means the user or App is new to the platform and has no rating history on it.

Outline

- Background
- Proposed Method
- Experiments and Results
- Conclusion

Dataset and Evaluation

Table I. Some Statistics of the iphone-iPad Dataset

	Amount	Min. #ratings	Max. #ratings	Avg. #ratings
User	112, 031	2	219	2.86
App-iPhone	2, 704	1	15, 946	62.31
App-iPad	2, 704	1	11, 846	56.23

We selected users who rated at least once on both of these platforms.

Table II. Some Statistics of the iphone-iPad-iMac Dataset

	Amount	Min. #ratings	Max. #ratings	Avg. #ratings
User	121, 905	2	21	2.21
App-iPhone	201	1	64, 482	1, 117.37
App-iPad	201	1	5, 572	206.62
App-iMac	201	1	439	13.97

We selected users who had at least two ratings on all platforms.

Rating Prediction

- MAE
- RMSE

Top-N Recommendation

- Recall
- NDCG

Research Questions

- **(RQ1).** How does STAR perform as compared to other state-of-the-art competitors?
- **(RQ2).** How is the performance of STAR in handling the new-user and new-App cold-start problems?
- **(RQ3).** Whether the rated App on current platform is the user's preferable one as compared to the same App on other unrated platforms?
- **(RQ4).** How do the common features and specific features of Apps contribute to the overall effectiveness of STAR?
- **(RQ5).** In addition to rating prediction that is prevalent to a recommendation algorithm, how does STAR perform in the more practical top-N recommendation?

Baseline Methods

- SVD++ [Koren 2008] (Collaborative Filtering)
- RMR [Ling et al. 2014] (Semantics Enhanced Recommendation)
- CTR [Wang and Blei 2011] (Semantics Enhanced Recommendation)
- FM [Rendle et al. 2011] (Context-Aware Recommender System)
- CMF [Singh and Gordon 2008] (Cross-Domain Recommender System)
- WMF [Hu et al. 2008] (Collaborative Filtering)
- Popular (Non-personalized method)

Overall Performance Comparisons (RQ1)

Table III. Performance comparison of various methods on the iPhone-iPad dataset regarding RMSE and MAE.

Methods	MAE	RMSE	p-value (MAE)	p-value (RMSE)
SVD++	0.7823 ± 0.003	1.1218 ± 0.004	$2.44e$-10	$1.02e$-10
RMR	0.7679 ± 0.002	1.0716 ± 0.003	$5.83e$-09	$7.16e$-08
FM	0.8002 ± 0.003	1.0887 ± 0.003	$3.05e$-11	$2.06e$-10
CMF	0.7701 ± 0.003	1.0977 ± 0.003	$2.96e$-09	$7.21e$-10
STAR	$\mathbf{0.7560 \pm 0.003}$	$\mathbf{1.0595 \pm 0.003}$	–	–

Table IV. Performance comparison of various methods on the iPhone-iPad-iMac dataset regarding RMSE and MAE..

Methods	MAE	RMSE	p-value (MAE)	p-value (RMSE)
SVD++	0.6040 ± 0.002	0.9389 ± 0.002	$2.88e$-09	$5.79e$-11
RMR	0.5985 ± 0.003	0.9333 ± 0.004	$1.77e$-08	$7.61e$-07
FM	0.6117 ± 0.003	0.9352 ± 0.003	$5.55e$-10	$2.19e$-08
CMF	0.6035 ± 0.003	0.9371 ± 0.003	$3.30e$-09	$6.60e$-10
STAR	$\mathbf{0.5889 \pm 0.002}$	$\mathbf{0.9261 \pm 0.003}$	–	–

Handling Cold-Start Problems (RQ2)

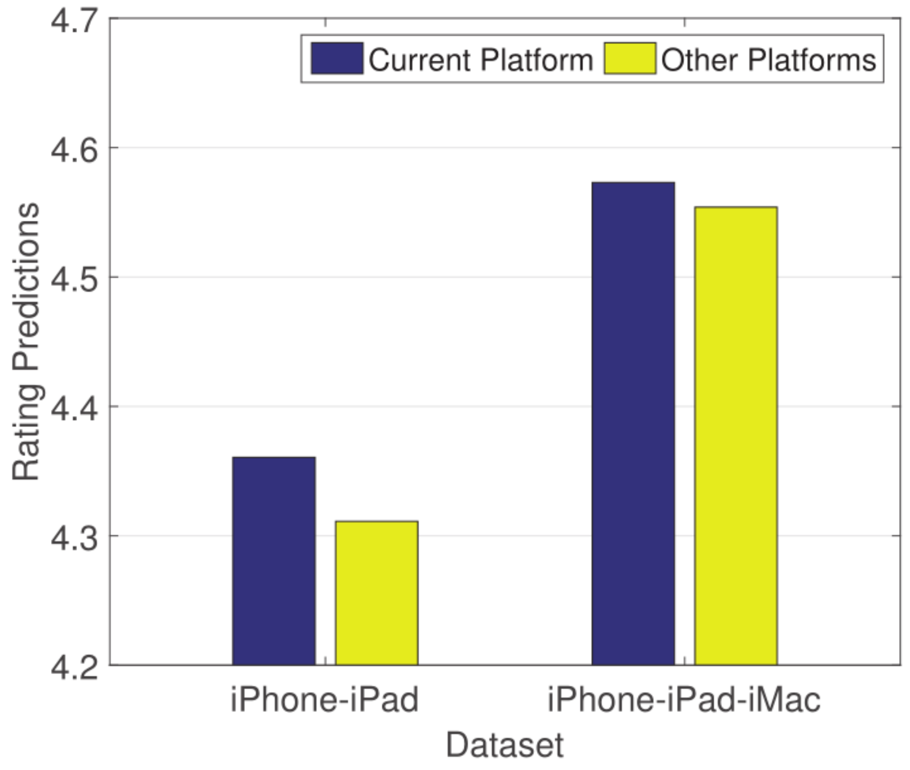
Table V. Performance comparison of various methods in handling new-user cold-start problem.

Methods	MAE	RMSE	p-value (MAE)	p-value (RMSE)
RMR	0.7003 ± 0.003	1.0592 ± 0.002	$2.67e$-09	$3.97e$-10
FM	0.7223 ± 0.003	1.0690 ± 0.003	$7.67e$-11	$1.05e$-10
CMF	0.7041 ± 0.004	1.0783 ± 0.003	$1.10e$-09	$4.04e$-11
STAR	$\mathbf{0.6849 \pm 0.003}$	$\mathbf{1.0344 \pm 0.003}$	–	–

Table VI. Performance comparison of various methods in handling new-App cold-start problem.

Methods	MAE	RMSE	p-value (MAE)	p-value (RMSE)
CTR	0.8746 ± 0.003	1.2310 ± 0.003	$6.76e$-10	$6.98e$-10
FM	0.9017 ± 0.003	1.2452 ± 0.004	$2.64e$-11	$1.76e$-11
CMF	0.8736 ± 0.002	1.2471 ± 0.002	$8.17e$-10	$2.48e$-11
STAR	$\mathbf{0.8529 \pm 0.002}$	$\mathbf{1.2063 \pm 0.003}$	–	–

User Preference on App-Platform (RQ3)



1. Our method favors the current platform better.
2. The gap between rating predictions of current platform and other platforms on the iPhone-iPad dataset is larger than that of the iPhone-iPad-iMac dataset.

Justification of Common Features and Specific Features (RQ4)

Table VII. Importance comparisons of common features and specific features on the iPhone-iPad dataset.

Methods	MAE	RMSE	p-value (MAE)	p-value (RMSE)
cSTAR	0.7771 ± 0.004	1.0824 ± 0.003	$7.57e$-10	$8.72e$-09
sSTAR	0.7842 ± 0.003	1.1042 ± 0.002	$2.37e$-10	$5.39e$-13
STAR	$\mathbf{0.7560 \pm 0.003}$	$\mathbf{1.0595 \pm 0.003}$	–	–

Table VIII. Importance comparisons of common features and specific features on the iPhone-iPad-iMac dataset.

Methods	MAE	RMSE	p-value (MAE)	p-value (RMSE)
cSTAR	0.6092 ± 0.003	0.9364 ± 0.003	$8.83e$-10	$2.13e$-07
sSTAR	0.6299 ± 0.003	0.9471 ± 0.002	$5.31e$-11	$1.24e$-09
STAR	$\mathbf{0.5889 \pm 0.002}$	$\mathbf{0.9261 \pm 0.003}$	–	–

Evaluation of Top-N Recommendation (RQ5)

Table IX. Evaluation of Top-N Recommendation on the iphone-iPad Dataset

Methods	Recall@100	NDCG@100	p-value(Recall)	p-value(NDCG)
Popular	0.5227 ± 0.000	0.1757 ± 0.000	$1.58e\text{-}11$	$1.48e\text{-}11$
WMF	0.5695 ± 0.003	0.2158 ± 0.002	$1.02e\text{-}10$	$2.12e\text{-}09$
CMF-WMF	0.5696 ± 0.002	0.2241 ± 0.003	$2.74e\text{-}08$	$3.66e\text{-}08$
STAR-WMF	$\mathbf{0.5782 \pm 0.002}$	$\mathbf{0.2321 \pm 0.003}$	–	–

Table X. Evaluation of Top-N Recommendation on the iphone-iPad-iMac Dataset

Methods	Recall@10	NDCG@10	p-value(Recall)	p-value(NDCG)
Popular	0.6133 ± 0.000	0.5005 ± 0.000	$6.21e\text{-}12$	$4.42e\text{-}11$
WMF	0.6686 ± 0.002	0.5321 ± 0.003	$3.12e\text{-}09$	$9.19e\text{-}09$
CMF-WMF	0.6745 ± 0.003	0.5379 ± 0.004	$2.38e\text{-}08$	$1.63e\text{-}07$
STAR-WMF	$\mathbf{0.6834 \pm 0.003}$	$\mathbf{0.5434 \pm 0.003}$	–	–

Outline

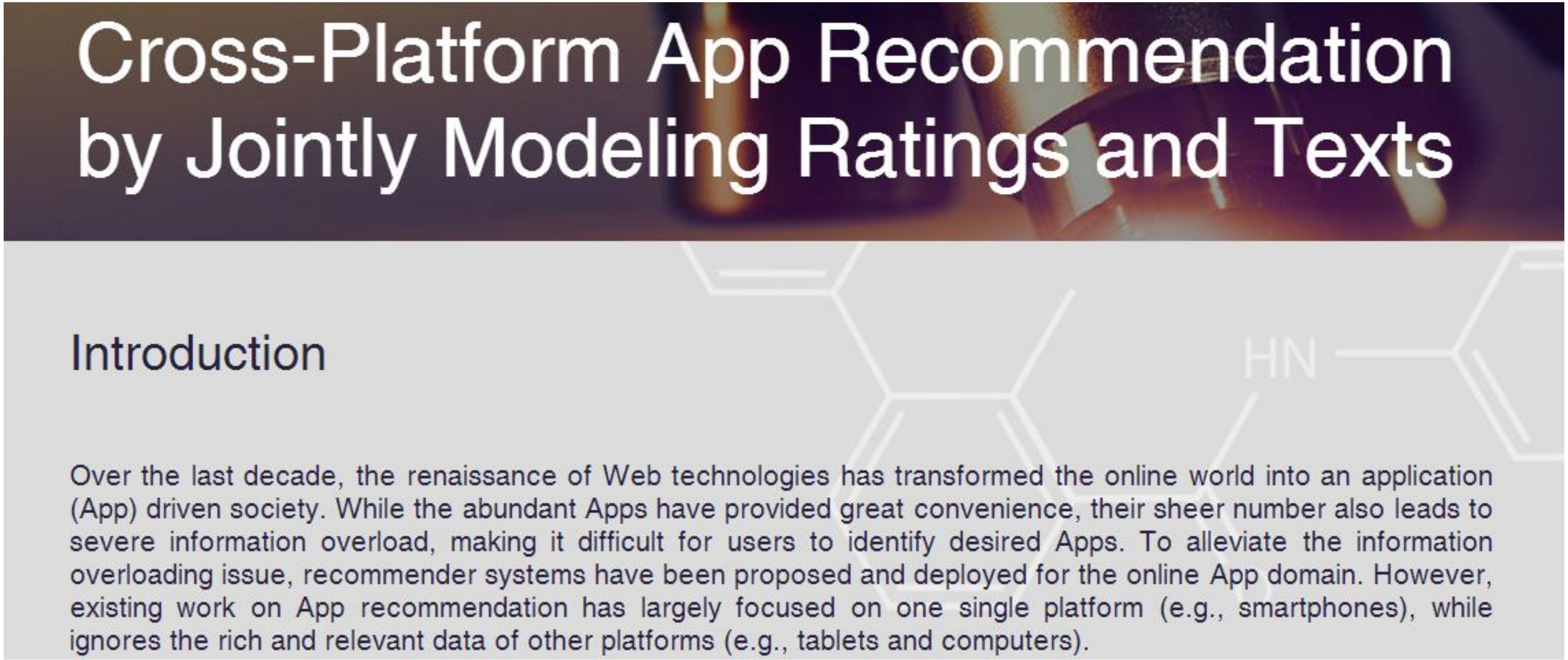
- Background
- Proposed Method
- Experiments and Results
- Conclusion

Challenges Solved

- Platform Variance
- Data Heterogeneity
- Data Sparsity
- Cold-Start Problem

Website

Data & code are available at http://apprec.wixsite.com/star



Cross-Platform App Recommendation by Jointly Modeling Ratings and Texts

Introduction

Over the last decade, the renaissance of Web technologies has transformed the online world into an application (App) driven society. While the abundant Apps have provided great convenience, their sheer number also leads to severe information overload, making it difficult for users to identify desired Apps. To alleviate the information overloading issue, recommender systems have been proposed and deployed for the online App domain. However, existing work on App recommendation has largely focused on one single platform (e.g., smartphones), while ignores the rich and relevant data of other platforms (e.g., tablets and computers).

Thanksgiving

Co-authors:

Xiangnan He
(NUS)

Liqiang Nie
(SDU)

Xiaochi Wei
(BIT)

Xia Hu
(Texax A&M)

Shunxiang Wu
(XMU)

Tat-Seng Chua
(NUS)

And all audiences…