Community-aware Social Recommendation: A Unified SCSVD Framework (Extended Abstract)

Jiewen Guan[†‡], Xin Huang[*] and Bilian Chen[†‡]
†Department of Automation, Xiamen University
‡Xiamen Key Laboratory of Big Data Intelligent Analysis and Decision-making
*Department of Computer Science, Hong Kong Baptist University

Motivations

Recommendation Systems

- Recommendation systems aim to accurately capture user preferences and provide high-quality recommendations.
- There is usually a network depicting socially-connected users in recommendation systems.

Community Structure

- A salient characteristic of social networks is the community structure inside them.
- Community structure has been shown to be informative for different downstream tasks, such as *link prediction*, etc.

A Natural Question: Will the community structure in social networks be helpful for improving recommendation?

Motivations

A Hypothesis Testing:

1. Conducting modularity-based community detection on the social network to get communities;
2. Computing the intra-community and inter-community user similarities of well-selected user pairs in terms of the Jaccard similarity, with all values saved in order as equal-length vectors $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ respectively;
3. Conducting a two-sample Student's t-test, where the null hypothesis is $H_0: \boldsymbol{\zeta} = \boldsymbol{\xi}$, while the alternative hypothesis is $H_1: \boldsymbol{\zeta} > \boldsymbol{\xi}$.

Experimental Results:

Dataset	Ciao	Epinions	Flixster
t-statistic	25.7776	27.0802	8.4342
p-value	2.0797e-142	5.7510e-161	1.7897e-17

Results of The t-Tests on The Three Datasets (will be introduced in experiments)

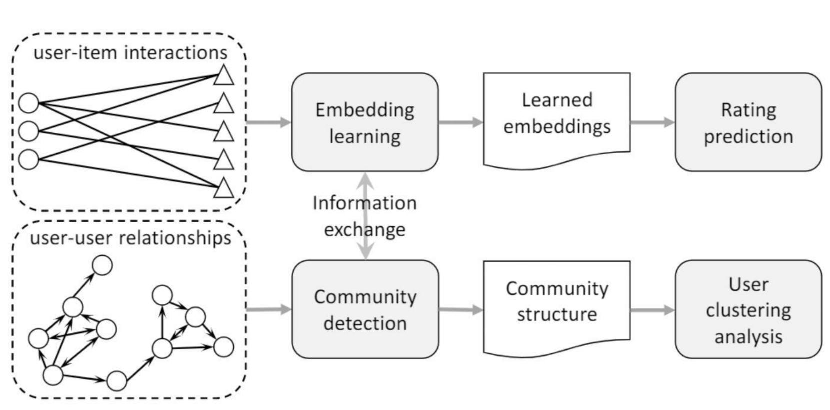
The SCSVD Framework

Philosophy: Conducting SVD and modularity-based community detection simultaneously, while building a bilateral connection between these two modules at the same time.

Optimization Model

$$
\begin{aligned}
\min_{\substack{\mathbf{p},\mathbf{q},\mathbf{U},\mathbf{V},\\ \mathbf{Y},\mathbf{C},\mathbf{H}}} \; & \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}\left(r_{ij}-\hat{r}_{ij}\right)^2 + \frac{\alpha}{2}\left\|\mathbf{H}-\mathbf{U}^T\mathbf{C}\right\|_F^2 \\
& -\frac{\beta}{2}\operatorname{Tr}\left(\mathbf{H}^T\left(\mathbf{A}\odot\mathbf{B}\right)\mathbf{H}\right) + \frac{\gamma}{2}(\|\mathbf{U}\boldsymbol{\Omega}_{\mathbf{U}}\|_F^2 \\
& + \|\mathbf{V}\boldsymbol{\Omega}_{\mathbf{V}}\|_F^2 + \|\mathbf{Y}\boldsymbol{\Omega}_{\mathbf{Y}}\|_F^2 + \|\boldsymbol{\Omega}_{\mathbf{U}}\mathbf{p}\|_2^2 + \|\boldsymbol{\Omega}_{\mathbf{V}}\mathbf{q}\|_2^2) \\
\text{s.t.} \quad & \mathbf{H}^T\mathbf{H} = \mathbf{I}, \mathbf{H} \geq 0,
\end{aligned}
$$

Schematic Overview



Theoretical Analysis

Computation Complexity of Optimization Algorithm: The overall computational complexity of our proposed optimization algorithm is $\mathcal{O}(\Phi(m^2nk + cn^2))$, where Φ, m, n, k, c are the number of iterations, items, users, latent factors and communities, respectively.

Functionary Mechanism of SCSVD:

- Users within the same detected community will be modeled with similar latent representations;
- Users with similar latent preferences will be partitioned into the same community.

In this way, the learned user preferences and the detected communities can exchange information, leading to a better user preference modeling.

Experiments

Datasets:

Statistics	Ciao	Epinions	Flixster	Statistics	Ciao	Epinions	Flixster
# of Users	7,375	40,163	10,000	# of Trustors	6,792	27,681	4,508
# of Items	105,114	139,738	19,342	# of Trustees	7,297	39,003	4,544
# of Ratings	284,086	664,823	457,882	# of Relations	111,781	442,979	12,144
Rating Density	0.037%	0.012%	0.237%	Relation Density	0.226%	0.041%	0.059%

Comparative Approaches:

Matrix Factorization-based Recommendation Models: CBSVD, SSLSVD, TrustSVD, SocialMF, SoRec, SoReg, PMF, SVD++ and TrustMF.

Neural Network-based Recommendation Models: NeuMF, GraphRec and DANSER.

Evaluation Metrics: MAE and RMSE.

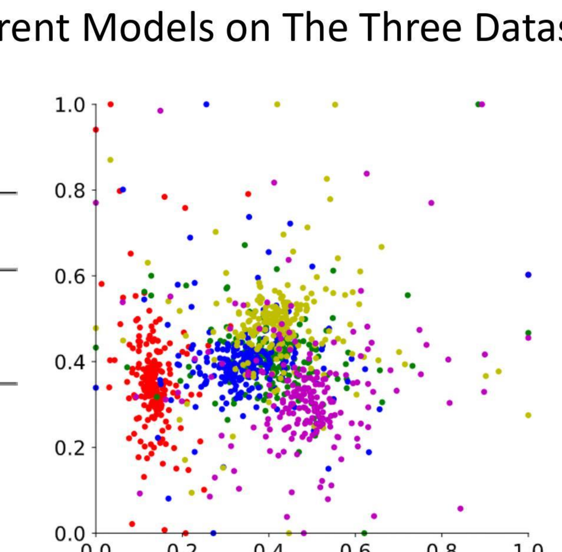
Experiments

Dataset		Metrics	CBSVD	SSLSVD	TrustSVD	SocialMF	SoRec	SoReg	PMF	SVD++	TrustMF	NeuMF	GraphRec	DANSER	SCSVD
Ciao	$k=5$	MAE	0.7214	0.7245	0.7253	0.7483	0.7649	0.7679	0.8861	0.7220	0.7470	0.7258	0.8171	0.7368	**0.7204**
		RMSE	0.9541	0.9570	0.9578	0.9795	0.9887	1.0382	1.1061	0.9581	0.9800	0.9716	1.0153	0.9659	**0.9530**
	$k=10$	MAE	0.7215	0.7252	0.7256	0.7553	0.7663	0.7659	0.8976	0.7229	0.7573	0.7251	0.7765	0.7463	**0.7209**
		RMSE	0.9543	0.9585	0.9589	0.9874	1.0124	1.0386	1.1140	0.9600	1.0125	0.9681	0.9893	0.9733	**0.9539**
Epinions	$k=5$	MAE	0.8023	0.8016	0.8016	0.8288	0.8384	0.8807	0.9120	0.8028	0.8200	0.7989	0.9337	0.8472	**0.7973**
		RMSE	1.0412	1.0455	1.0448	1.0810	1.0738	1.1564	1.1805	1.0487	1.0716	1.0480	1.1293	1.0846	**1.0402**
	$k=10$	MAE	0.8023	0.8023	0.8025	0.8375	0.8374	0.8480	0.9431	0.8052	0.8294	**0.7951**	0.8534	0.8472	0.7974
		RMSE	1.0412	1.0466	1.0460	1.0961	1.1022	1.1183	1.2107	1.0516	1.1053	1.0493	1.0798	1.0847	**1.0408**
Flixster	$k=5$	MAE	0.6662	0.6649	0.6642	0.6853	0.6961	0.7115	0.6995	0.6670	0.6869	0.6701	0.7933	0.6899	**0.6639**
		RMSE	0.8899	0.8864	0.8854	0.9081	0.9253	0.9446	0.9338	0.8926	0.9073	0.8914	0.9952	0.9185	**0.8839**
	$k=10$	MAE	0.6641	0.6638	0.6632	0.6876	0.7037	0.6977	0.7013	0.6659	0.6893	0.6655	0.7349	0.6930	**0.6615**
		RMSE	0.8878	0.8865	0.8853	0.9139	0.9390	0.9290	0.9368	0.8929	0.9134	0.8879	0.9492	0.9217	**0.8815**

Rating Prediction Performance Comparison of Different Models on The Three Datasets

Dataset		SVD++	SCSVD	DANSER	GraphRec
Name	# Users				
Ciao	7,375	183.95	229.28	1267.38	4885.23
Flixster	10,000	1091.74	1297.74	5986.37	13411.63
Epinions	40,163	196.67	279.82	2353.07	14098.38

Running Time Comparison (in second/epoch)



User Embedding Visualization (different colors represent users in different communities)