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

Jiewen Guan*†, Xin Huang‡, Bilian Chen*†

*Department of Automation, Xiamen University, Xiamen 361005, China

†Xiamen Key Laboratory of Big Data Intelligent Analysis and Decision-making, Xiamen 361005, China

‡Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China

Email: jwguan@stu.xmu.edu.cn, xinhuang@comp.hkbu.edu.hk, blchen@xmu.edu.cn

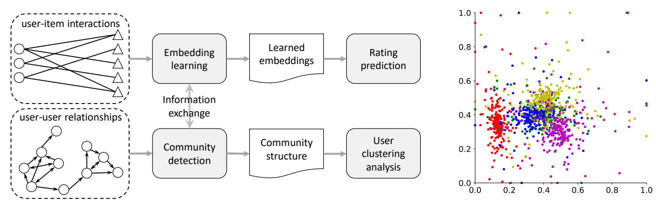
***Abstract*—Social recommendation aims at improving recommendation performance by incorporating social information. Most existing social recommender systems only utilize the one-hop interpersonal social information, neglecting the community structure emerged in social networks, which may contain additional conducive information. In this paper, we propose a unified Simultaneous Community detection and Singular Value Decomposition (SCSVD) framework for community-aware social recommendation. An efficient optimization algorithm is also derived to optimize SCSVD, with an analysis of convergence and computational complexity. Comprehensive experimental results on three real-world benchmark datasets demonstrate the effectiveness of SCSVD, over both traditional matrix factorization based recommendation models and advanced neural network based recommendation models.**

I. Introduction

With the rapid development of online social media, people can nowadays issue a variety of social activities in shopping platforms. According to social homophily and social influence theories [1], users with more social interactions are more likely to share similar preferences, leading to *social recommendation*. Social recommendation aims at augmenting recommendation performance of traditional recommender systems by taking additional social information into account, which has achieved a great success during the last decade. Most existing social recommender systems only consider the one-hop social connections between users, but ignore multi-hop connected communities.

A salient characteristic of social networks is the community structure inside them [2]. In a social network, users prefer to interact more frequently with other members within the same community. Community structure has been shown to be informative for different downstream tasks, such as *link prediction* [3], etc. Naturally, a question turns out: Will the community structure in social networks be helpful for improving recommendation?

In this paper, we first statistically show that communities in social networks are really conducive for improving recommendation. Afterwards, we propose a unified Simultaneous Community detection and Singular Value Decomposition (SCSVD) framework for community-aware social recommendation, which builds on top of SVD++ [5] and modularity-based community detection [4]. To optimize SCSVD, we derive an effective optimization algorithm. We theoretically analyze the computational complexity, convergence, and functionary mechanism of SCSVD. Fig. 1a shows an overview of SCSVD.



(a) An overview of the SCSVD framework (b) Visualization of $\mathbf{U}$

Fig. 1: An overview of the SCSVD framework and the embedding visualization on the Epinions dataset.

II. Motivation: A Hypothesis Testing

In this section, we present a hypothesis testing to answer the following question from a statistical perspective to motivate the idea of community-aware social recommendation: Do users belonging to the same community have more similar rating patterns?

Specifically, we conduct the following three-stage test on three real-world benchmark product review datasets (Ciao, Epinions and Flixster) to answer the above question: 1) Conducting modularity-based community detection [4] 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 $H_1 : \boldsymbol{\zeta} > \boldsymbol{\xi}$.

TABLE I: Results of The t-Tests on The Three Datasets

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

The t-test results in Table I show that the null hypothesis H_0 is strongly rejected with very high confidence, which affirmatively answers the aforementioned question and motivates the idea of community-aware social recommendation.

TABLE II: Rating Prediction Performance Comparison of Different Models on The Three Datasets

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**

III. Summary of SCSVD

The SCSVD framework [6]. The SCSVD seeks to conduct SVD++ [5] and modularity-based community detection [4] simultaneously, while building a bilateral connection between these two modules at the same time. In this way, the underlying community structure in social networks can be exploited to guide the process of user latent preference modeling. The optimization problem of our SCSVD model is formulated as

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

where the symbol definitions can be found in our full paper [6].

To optimize SCSVD, we propose an effective optimization algorithm in the following computational complexity.

Theorem 1. *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.*

The functionary mechanism of SCSVD. The functionary mechanism of SCSVD appears as a bilateral information exchange process: 1) Users within the same detected community will be modeled with similar latent representations; 2) Users with similar latent preferences will be partitioned into the same community. By this mechanism, the learned user preferences from rating information and the detected communities from social connections can exchange information with each other, leading to a more precise user preference modeling.

IV. Experimental Study

In this section, we show some sample experimental results. More experimental results and more details about the experimental settings can be found in our full paper [6].

Experimental settings. 1) **Datasets**: We adopt three real-world benchmark product review datasets: Ciao, Epinions and Flixster. 2) **Comparative methods**: We compare SCSVD against twelve state-of-the-art methods: CBSVD, SSLSVD, TrustSVD, SocialMF, SoRec, SoReg, PMF, SVD++, TrustMF, NeuMF, GraphRec and DANSER. 3) **Evaluation metrics**: We adopt MAE and RMSE as rating prediction performance indicators.

Exp-1: Rating prediction performance comparison. The experimental results of rating prediction of all methods on the three datasets are presented in Table II. As observed, our proposed SCSVD method performs best in all except for one case, which illustrates its powerful and stable performance.

Exp-2: User embedding visualization. We perform dimensionality reduction on the learned user latent preferences (denoted by $\mathbf{U}$) via the PCA tool. We plot the low-dimensional user embeddings in $\mathbb{R}^2$ as shown in Fig. 1b, where different colors represent different detected communities. As observed, the learned user embeddings manifest a relatively clear community structure, which is in accord with our analysis.

Exp-3: Efficiency evaluation. We test the efficiency of SCSVD. Specifically, we run SVD++, SCSVD, DANSER and GraphRec on the three datasets. We report their training time in Table III. As observed, SCSVD is slightly slower than SVD++, but evidently faster than DANSER and GraphRec, which shows its efficiency to some extent.

TABLE III: Running Time Comparison (in second/epoch)

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

Acknowledgment

The work was supported in part by Youth Innovation Fund of Xiamen (Grant No. 3502Z20206049), National Natural Science Foundation of China (Grant No. 61836005) and HK RGC (Grant No. 22200320). Bilian Chen is the corresponding author.

References

[1] R. Zafarani, M. A. Abbasi, and H. Liu, *Social Media Mining: An Introduction*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[2] X. Huang, L. V. Lakshmanan, and J. Xu, *Community Search over Big Graphs*. San Rafael, CA, USA: Morgan & Claypool Pub., 2019.

[3] J. Tang, H. Gao, X. Hu, and H. Liu, "Exploiting homophily effect for trust prediction", in *Proc. 6th ACM Int. Conf. Web Search Data Mining*, 2013, pp. 53-62.

[4] E. A. Leicht and M. E. J. Newman, "Community structure in directed networks", *Phy. Rev. Lett.*, vol. 100, no. 11, 2008, Art. no. 118703.

[5] Y. Koren, "Factorization meets the neighborhood: A multifaceted collaborative filtering model", in *Proc. 14th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2008, pp. 426-434.

[6] J. Guan, X. Huang, B. Chen, "Community-aware social recommendation: A unified SCSVD framework", *IEEE Trans. Knowl. Data Eng.*, Early Access, 2021.