35th AAAI Conference on Artificial Intelligence
A virtual Conference, February 2-9, 2021

Semi-Discrete Social Recommendation (Student Abstract)

Fangyuan Luo[1], Jun Wu[1], HaishuaiWang[2]
[1]Beijing Jiaotong University, Beijing [2]Harvard University, Cambridge

BEIJING JIAOTONG UNIVERSITY

Motivation

Contributions

1、We combines MF model with NE model partially under the discrete constraints, which is essential in modeling the influence propagation process of users' interests over social network, as well as avoiding the extra encoding loss caused by the discretization on social embedding.

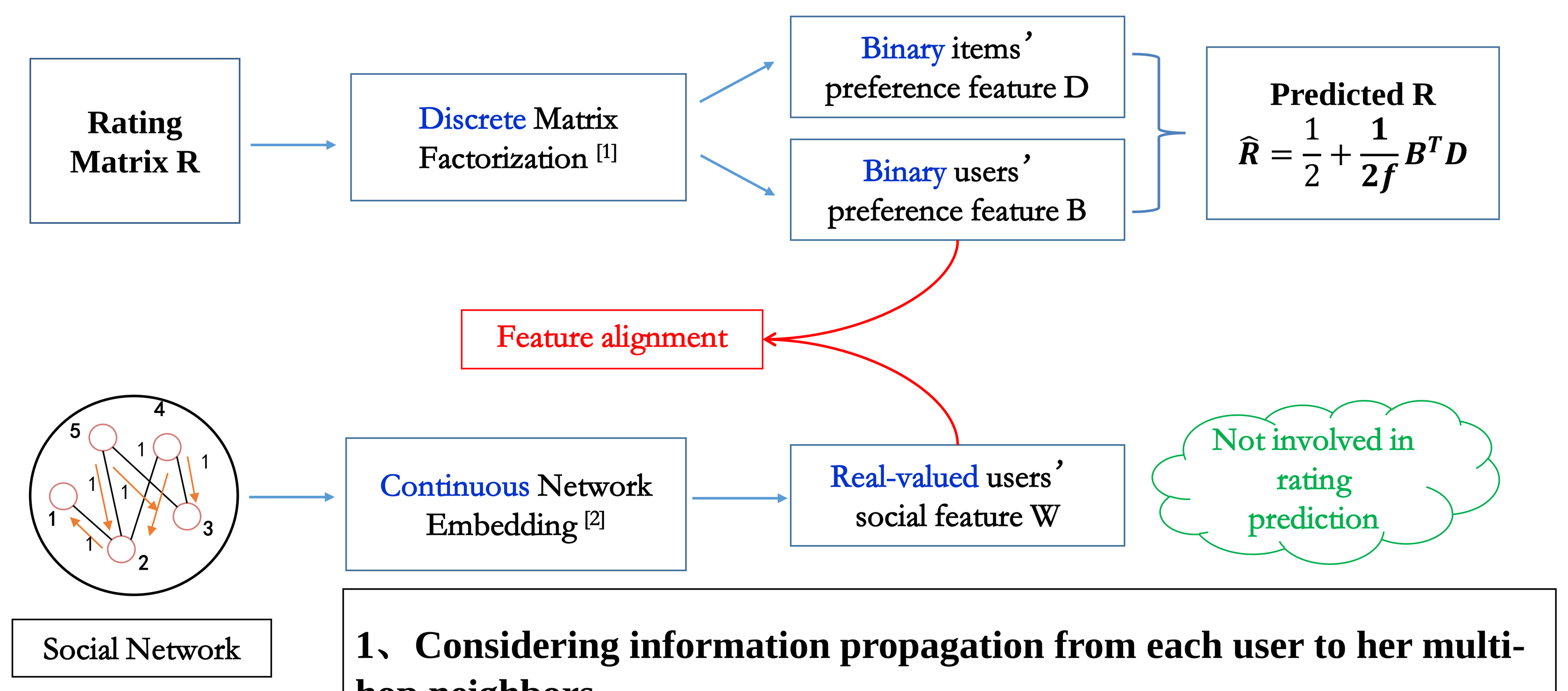
2、We develop an efficient alternating optimization algorithm to solve the proposed mixed-integer programming problem, which can yield informative and compact hash codes.

References

[1] Zhang, H.; Shen, F.; Liu, W.; He, X.; Luan, H.; and Chua, T. S. 2016. Discrete Collaborative Filtering. In SIGIR, 325–334.
[2] Perozzi, B.; Alrfou, R.; and Skiena, S. 2014. DeepWalk: online learning of social representations. In KDD, 701–710.
[3] F. Shen, C. Shen, W. Liu, H. T. Shen, Supervised discrete hashing, in: CVPR, 2015, pp. 37-45.

Methods

Framework:



1、Considering information propagation from each user to her multi-hop neighbors
2、 Avoiding additional encoding loss from unnecessary discretization

Formulation:

$$\underset{\mathbf{B},\mathbf{D},\mathbf{W},\mathbf{V}}{\arg\min} \boxed{\sum_{(u,i)\in\Omega} (r_{ui} - s_H(\mathbf{b}_u, \mathbf{d}_i))^2} + \boxed{\alpha \sum_u ||\mathbf{b}_u - \mathbf{w}_u||^2}$$
$$- \beta \boxed{\sum_u [\log \sigma(\mathbf{w}_u^T \mathbf{h}) + \log \sigma(-\mathbf{w}_u^T \bar{\mathbf{h}})]}$$
$$+ \frac{\lambda}{2}(||\mathbf{W}||_F^2 + ||\mathbf{V}||_F^2)$$
$$s.t. \quad \boxed{\begin{array}{l}\mathbf{B} \in \{\pm 1\}^{f\times m}, \mathbf{D} \in \{\pm 1\}^{f\times n}, \\ \mathbf{B}\mathbf{1}_m = 0, \mathbf{D}\mathbf{1}_n = 0, \mathbf{B}\mathbf{B}^T = m\mathbf{I}_f, \mathbf{D}\mathbf{D}^T = n\mathbf{I}_f.\end{array}}$$

Square Loss

Feature Alignment

Network Embedding Loss

Discrete, Balanced and Decorrelation Constraints

Optimization

NP-Hard➡Soften B and D:

$$\mathbf{X} \in \mathcal{B} = \{\mathbf{X} \in \mathbb{R}^{f\times m} | \mathbf{X}\mathbf{1} = 0, \mathbf{X}\mathbf{X}^{\mathbf{T}} = m\mathbf{I}\} \qquad d(\mathbf{B}, \mathcal{B}) = \min_{\mathbf{X}\in\mathcal{B}} ||\mathbf{B} - \mathbf{X}||_F$$
$$\mathbf{Y} \in \mathcal{D} = \{\mathbf{Y} \in \mathbb{R}^{f\times n} | \mathbf{Y}\mathbf{1} = 0, \mathbf{Y}\mathbf{Y}^{\mathbf{T}} = n\mathbf{I}\} \qquad d(\mathbf{D}, \mathcal{D}) = \min_{\mathbf{Y}\in\mathcal{D}} ||\mathbf{D} - \mathbf{Y}||_F$$

$$\underset{\mathbf{B},\mathbf{D},\mathbf{W},\mathbf{V}}{\arg\min} \sum_{(u,i)\in\Omega} (r_{ui} - s_H(\mathbf{b}_u, \mathbf{d}_i))^2 - \alpha \sum_u [\log \sigma(\mathbf{w}_u^T \mathbf{h}) + \log \sigma(-\mathbf{w}_u^T \bar{\mathbf{h}})]$$
$$+ \beta \sum_u ||\mathbf{b}_u - \mathbf{w}_u||^2 + \frac{\lambda_w}{2}||\mathbf{W}||_F^2 + \frac{\lambda_v}{2}||\mathbf{V}||_F^2$$
$$+ \lambda_B d^2(\mathbf{B}, \mathcal{B}) + \lambda_D d^2(\mathbf{D}, \mathcal{D})$$
$$s.t. \quad \mathbf{B} \in \{\pm 1\}^{f\times m}, \mathbf{D} \in \{\pm 1\}^{f\times n},$$

B-subproblem:

$$\underset{\mathbf{b}_u \in \{\pm 1\}^f}{\arg\min} \sum_{i\in\Omega_u} \frac{1}{4f^2}(\mathbf{d}_i^T \mathbf{b}_u)^2 - \frac{1}{f}(r_{ui} - \frac{1}{2})\mathbf{d}_i^T \mathbf{b}_u - 2\beta \mathbf{w}_u^T \mathbf{b}_u - 2\lambda_B \mathbf{x}_u^T \mathbf{b}_u$$

DCD [3]

$$\underset{b_{uk} \in \{\pm 1\}}{\arg\min} \hat{b}_{uk} b_{uk} \qquad \boxed{\hat{b}_{uk} = \sum_{i\in\Omega_u} \frac{1}{f}(r_{ui} - \frac{1}{2} - \frac{1}{2f}\mathbf{d}_i^T \mathbf{b}_u) d_{ik} + \frac{1}{2f^2} b_{uk} + 2\beta w_{uk} + 2\lambda_B x_{uk}}$$

D-subproblem:

$$\underset{\mathbf{d}_i \in \{\pm 1\}^f}{\arg\min} \sum_{u\in\Omega_i} \frac{1}{4f^2}(\mathbf{b}_u^T \mathbf{d}_i)^2 - \frac{1}{f}(r_{ui} - \frac{1}{2})\mathbf{b}_u^T \mathbf{d}_i - 2\lambda_D \mathbf{y}_i^T \mathbf{d}_i$$

DCD

$$\underset{d_{ik} \in \{\pm 1\}}{\arg\min} \hat{d}_{ik} d_{ik} \qquad \boxed{\hat{d}_{ik} = \sum_{u\in\Omega_i} \frac{1}{f}(r_{ui} - \frac{1}{2} - \frac{1}{2f}\mathbf{b}_u^T \mathbf{d}_i) b_{uk} + \frac{1}{2f^2} d_{ik} + 2\lambda_D y_{ik}}$$

W & V-subproblems: SGD + BP

X & Y-subproblems: SVD + Gram-Schmidt orthogonalization

Experiment

Datasets:

Datasets	Users	Items	Ratings	Rating Density	Followers	Followees	Links	Linking Density
CiaoDVD	17,615	16,121	72,665	0.026%	1,438	4,299	40,133	0.649%
Epinions	40,163	139,738	664,824	0.012%	33,960	49,288	487,183	0.029%

- S2MF, DSR/DTMF > DCF: the effectiveness of social information to hash code learning
- S2MF > DSR/DTMF: the superiority of our solution to social information processing
- SMF > S2MF > MF > DSR/DTMF > DCF: the effectiveness of our proposed method

Results:

		CiaoDVD				Epinions			
Methods	Metric	8 Bits	16 Bits	32 Bits	64 Bits	8 Bits	16 Bits	32 Bits	64 Bits
SMF	NDCG@10	0.82807	0.83035	0.82818	0.83171	0.84666	0.84680	0.84639	0.84801
	Improve	-1.61%	-1.28%	-1.48%	-1.27%	-3.50%	-2.26%	-1.87%	-1.77%
MF	NDCG@10	0.75201	0.76588	0.76055	0.76881	0.79658	0.80503	0.80629	0.80686
	Improve	+8.34%	+7.03%	+7.28%	+6.81%	+2.57%	+2.81%	+3.01%	+3.24%
DSR/DTMF	NDCG@10	0.74898	0.75168	0.75942	0.76142	0.78722	0.78880	0.79586	0.79665
	Improve	+8.78%	+9.05%	+7.44%	+7.84%	+3.79%	+4.93%	+4.36%	+4.57%
DCF	NDCG@10	0.72930	0.73269	0.74305	0.74921	0.77929	0.78257	0.78837	0.79123
	Improve	+11.72%	+11.87%	+9.81%	+9.60%	+4.84%	+5.77%	+5.35%	+5.28%
S2MF	NDCG@10	**0.81475**	**0.81969**	**0.81595**	**0.82114**	**0.81702**	**0.82769**	**0.83053**	**0.83302**