SCIENCE CHINA
Information Sciences

• LETTER •

April 2022, Vol. 65 149104:1–149104:3
https://doi.org/10.1007/s11432-019-2884-0

Leveraging implicit social structures for recommendation via a Bayesian generative model

Huafeng LIU[1,2], Jingxuan WEN[1,2], Liping JING[1,2*] & Jian YU[1,2]

[1]*School of Computer and Information Technology, Beijing Jiaotong University, Beijing* 100044*, China;*
[2]*Beijing Key Lab of Traffic Data Analysis and Mining, Beijing Jiaotong University, Beijing* 100044*, China*

Received 11 August 2019/Revised 28 October 2019/Accepted 11 December 2019/Published online 3 March 2022

Citation Liu H F, Wen J X, Jing L P, et al. Leveraging implicit social structures for recommendation via a Bayesian generative model. Sci China Inf Sci, 2022, 65(4): 149104, https://doi.org/10.1007/s11432-019-2884-0

Dear editor,

Social recommendation leverages the social relations between users and their past behaviors to model user preferences. While social recommendation is a core component in recommendation systems, discovering and utilizing hidden patterns among complex social networks and user behavior data is not a trivial problem.

This study focuses on the cold-start problem of recommendation systems and attempts to solve it with the aid of social relations. Thus, we propose a new matrix approximation model by leveraging implicit social structures for recommendation (SoMA). SoMA takes advantage of social networks and user preference information via a joint graphical model to determine latent user and item factors. Specifically, any pairwise user social relation (including both observed and unknown social relations) is modeled by considering both local and global social structures. Among them, local social structures are characterized by user-overlapping groups, while the global structure is characterized by a user's popularity across the entire social network. Moreover, the user-item rating is modeled by latent user and item factors with the aid of discrete rating distribution. In this case, the latent user factors are shared by these two generative procedures, one for social relationships and the other for rating data, so that two generative procedures can affect each other. In this case, SoMA can be considered as a unified process that makes social information and preference information to sufficiently cooperate each other. Thus, SoMA has the ability to simultaneously and seamlessly capture the user's social properties and personal preferences for items and is expected to improve the recommendation quality.

The proposed model. In social networks, the relationships between two users are usually affected by two main factors: one is the users' own property, and the other is the social group (community) to which the user belongs. For example, if a user is popular, he or she will be more likely to establish relations with others. If two users belong to the same social group, the extent to which they will have a stronger social relation become larger. Thus, we consider these two characteristics to model user pairwise social relation.

To characterize user popularity property, we introduce a variable γ_i to indicate the user's popularity in the whole social network, which can be sampled from the following exponential distribution with rate parameter λ:

$$\boldsymbol{\gamma}_i \sim \mathrm{Exp}(\lambda). \tag{1}$$

Because of the diversity of users' interests, each user may belong to one or more social groups. To capture such types of local social structures, we introduce a user-group membership vector for each user in all social groups, $\boldsymbol{\theta}_i = \{\boldsymbol{\theta}_{ic}\}_{c=1}^{d}$, where d is the total number of groups. Among it, $\boldsymbol{\theta}_{ic} \in [0, 1]$ is the affiliation strength that user i belongs to social group c, which can be sampled from Beta distributions with the parameters α_1 and α_2 as follows:

$$\boldsymbol{\theta}_{ic} \sim \mathrm{Beta}(\alpha_1, \alpha_2). \tag{2}$$

To model the pairwise social relation among users, a variable $\boldsymbol{z}_{ikc}$ is introduced to indicate whether the i-th user belongs to group c when forming a relation with user k. As an indicator, $\boldsymbol{z}_{ikc}$ can be drawn from a Bernoulli distribution parameterized by the user-group membership $\boldsymbol{\theta}_{ic}$:

$$\boldsymbol{z}_{ikc} \sim B(\boldsymbol{\theta}_{ic}), \tag{3}$$

where Bernoulli distribution is adopted because it is conjugate with Beta distribution. Note that for directed social relations, $\boldsymbol{z}_{kic}$ is drawn by $\boldsymbol{z}_{kic} \sim B(\boldsymbol{\theta}_{kc})$, which is different from $\boldsymbol{z}_{ikc}$.

Because one user may be assigned to several social groups, his or her social relations will be affected by the corresponding social groups. To evaluate the contribution of each social group, we introduce a variable $\boldsymbol{\pi}_c$ that represents the strength of connectivity in group c as follows:

$$\boldsymbol{\pi}_c \sim \mathrm{Beta}(\beta_1, \beta_2), \tag{4}$$

where $\boldsymbol{\pi}_c \in [0, 1]$. Because the social groups may share common users, the social strength between two users may be generated by several groups. Naturally, the more groups that these users share, the larger is the extent to which these

* Corresponding author (email: lpjing@bjtu.edu.cn)

users are connected, i.e., their social strength will be considerably stronger. To model this property, a variable $\boldsymbol{q}_{ikc}$ is used to denote the strength of the social relationship $\langle i, k\rangle$ formed by the c-th group:

$$\boldsymbol{q}_{ikc} = \begin{cases} \boldsymbol{\pi}_c, & \text{if } \boldsymbol{z}_{ikc} = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Then, the probability ($\boldsymbol{p}_{ik}$) that the social relation $\langle i, k\rangle$ exists can be formulated as

$$\boldsymbol{p}_{ik} = 1 - \prod_{c=1}^{d}(1 - \boldsymbol{q}_{ikc}) = 1 - \prod_{c=1}^{d}(1 - \boldsymbol{\pi}_c)^{\boldsymbol{z}_{ikc}}. \tag{6}$$

As we know, the social relation may be directed. In this case, the probability ($\boldsymbol{p}_{ik}$) that a social relation $\langle i, k\rangle$ exists is different from the probability ($\boldsymbol{p}_{ki}$) that can be calculated by $\boldsymbol{p}_{ki} = 1 - \prod_{c=1}^{d}(1 - \boldsymbol{\pi}_c)^{\boldsymbol{z}_{kic}}$.

By combining the global social structure (i.e. the user's popularity (1)) and the local social structure (i.e., the user's social group (6)), we can model any pairwise social relation as follows:

$$\boldsymbol{S}_{ik} \sim B(g(\boldsymbol{\gamma}_i + \boldsymbol{\gamma}_k + \boldsymbol{p}_{ik})). \tag{7}$$

As the mean of the Bernoulli distribution is in $[0, 1]$, to learn parameters in a proper manner, we map $(\boldsymbol{\gamma}_i + \boldsymbol{\gamma}_k + \boldsymbol{p}_{ik})$ into $[0, 1]$ via the sigmoid function $g(x) = 1/(1 + \exp(-x))$.

Thus far, the latent space with d features captures the social information, especially the overlapping social groups, and each latent feature corresponds to one social group. Meanwhile, each user is modeled by a vector $\boldsymbol{\theta}_i \in \mathbb{R}^d$ in this latent space. Recall that our goal is to predict the correlation between users and items, and each item is modeled via a vector $\boldsymbol{\phi}_j \in \mathbb{R}^d$ that can be sampled by a Gaussian distribution with mean 0 and variance (σ_ϕ^2) as $\boldsymbol{\phi}_j \sim N(0, \sigma_\phi^2)$.

Considering each observed rating $\boldsymbol{R}_{ij}$ is discrete, a discrete distribution should be introduced [1]. A principled approach to modeling discrete data is to employ a properly normalized exponential family model. Denoting the preference data set by $\mathcal{Y} = \{k\}_{k=0}^{\mathcal{B}}$, where $\mathcal{B}$ is the maximum value of the preference data (e.g., the maximum rating score or the maximum number of clicks in recommendation systems). To normalize the Gaussian distribution, we use $\log\sum_{k=0}^{\mathcal{B}}\exp(-\frac{1}{2\sigma^2}(k-\mu)^2)$ to replace the original normalizer $\log(\sigma\sqrt{2\pi})$, and then

$$p(x|\mu, \sigma^2) = \exp\left(t(x)^{\mathrm{T}}\eta(\mu, \sigma) - \log\sum_{k=0}^{\mathcal{B}}\exp(t(k)^{\mathrm{T}}\eta(\mu, \sigma))\right),$$

where $t(x) = [(x - \frac{\mathcal{B}}{2}), (x - \frac{\mathcal{B}}{2})^2]^{\mathrm{T}}$, and $\eta(\mu, \sigma) = [\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2}]^{\mathrm{T}}$.

Thus far, the above distribution can characterize the discrete preference data in a discrete finite domain. To introduce the influence mechanism without additional variables, we define $\mu = \boldsymbol{\theta}_i^{\mathrm{T}}\boldsymbol{\phi}_j$.

In this case, the discrete preference value $\boldsymbol{R}_{ij}$ can be generated via discrete Gaussian distribution $N_d(\boldsymbol{R}_{ij}|\boldsymbol{\theta}_i, \boldsymbol{\phi}_j, \sigma_R^2)$ defined as

$$p(\boldsymbol{R}_{ij}|\boldsymbol{\theta}_i, \boldsymbol{\phi}_j, \sigma_R^2) = \exp\left(t(\boldsymbol{R}_{ij})^{\mathrm{T}}\eta(\boldsymbol{\theta}, \boldsymbol{\phi}) - A(\boldsymbol{\theta}, \boldsymbol{\phi})\right),$$

where $t(\boldsymbol{R}_{ij}) = [(\boldsymbol{R}_{ij} - \mathcal{B}/2), (\boldsymbol{R}_{ij} - \mathcal{B}/2)^2]^{\mathrm{T}}$, $\eta(\boldsymbol{\theta}, \boldsymbol{\phi}) = [\frac{\boldsymbol{\theta}_i^{\mathrm{T}}\boldsymbol{\phi}_j}{\sigma_R^2}, -\frac{1}{2\sigma_R^2}]^{\mathrm{T}}$, and $A(\boldsymbol{\theta}, \boldsymbol{\phi}) = \log\sum_{k=0}^{\mathcal{B}}\exp(t(k)^{\mathrm{T}}\eta(\boldsymbol{\theta}, \boldsymbol{\phi}))$. We theoretically analyze the generalization error bounds to guarantee model performance, see Appendix A.

Based on the generative process, all variables can be optimized by maximizing the posterior of all latent variables. According to Bayesian rule, it can be implemented by maximizing the joint probability of preference data, social data and all latent variables. Following the generative process of SoMA model, the joint probability can be expressed as

$$\begin{aligned} p(\Theta, \boldsymbol{R}, \boldsymbol{S}|\alpha, \beta, \lambda, \sigma_\phi, \sigma_R) &= \prod_{\boldsymbol{R}_{ij}\neq 0} p(\boldsymbol{R}_{ij}|\boldsymbol{\theta}_i, \boldsymbol{\phi}_j, \sigma_R) \\ &\times \prod_{\boldsymbol{S}_{ik}} p(\boldsymbol{S}_{ik}|\boldsymbol{z}_{ik}, \boldsymbol{\pi}) \prod_{i=1}^{n} p(\boldsymbol{\gamma}_i|\lambda) \prod_{i=1}^{n}\prod_{c=1}^{d} p(\boldsymbol{\theta}_{ic}|\alpha) \\ &\times \prod_{j=1}^{m} p(\boldsymbol{\phi}_j|\sigma_\phi) \prod_{\langle i,k\rangle}\prod_{c=1}^{d} p(\boldsymbol{z}_{ikc}|\boldsymbol{\theta}_{ic}) \prod_{c=1}^{d} p(\boldsymbol{\pi}_c|\beta), \end{aligned} \tag{8}$$

where $\boldsymbol{\theta}$ indicates the model parameters set. The key problem in the parameter estimation of Bayesian models is computing the posterior distribution over the latent variables given the observed data. For SoMA, obtaining the exact inference of the posterior distribution is intractable; therefore, we adopt the mean-field stochastic variational Bayesian method to approximate the full posterior distribution, and a detailed inference algorithm can be obtained, which is provided in Appendix B.

Experiments. We report the experimental results on the Epinions[1)] and Yelp[2)] datasets that comprise user item interactions and social relations among users. A series of experiments are conducted to compare the proposed SoMA with several baselines on All Users (all ratings are used as the testing set) and Near-cold-start Users (the users who rate less than five items will be involved in the testing set), as shown in Table 1 [2–4]. Obviously, SoMA performs better than the existing social recommendation methods, thereby confirming that consideration of both observed and unknown social relations is helpful to extract social structure.

Obviously, SoMA performs better than the existing social recommendation methods, thereby confirming that consideration of both observed and unknown social relations is helpful to extract social structure and leverage training of the social recommendation model. Even though SoDimRec makes use of observed and unobserved social relations, it adopts a two-phase separation strategy that ignores the interaction between social structure identification and recommendation model training. Note that deep social recommendation methods (NSCR DeepSoR) also achieve comparable performance because of non-linear feature learning. SoMA performs well because it has the ability to capture hidden social structures and improve recommendation performance in a unified pattern. More detailed experimental results have been provided in Appendix C.

Conclusion. This study proposed a new social recommendation model (SoMA). SoMA takes advantage of social overlapping user group detecting on social networks and matrix factorization collaborative filtering of user behavior information to learn latent user and item factors. SoMA is good at exploiting social relations and rating information. Thus,

1) http://www.trustlet.org/downloaded_epinions.html.
2) https://www.yelp.com/dataset/challenge.

Table 1 Comparing different recommendation methods for testing All Users and Near-cold-start Users

Method	All Users				Near-cold-start Users			
	Epinions		Yelp		Epinions		Yelp	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
SoDimRec [2]	0.9305	1.1524	0.6537	0.8296	1.0689	1.2606	0.7461	0.9194
NSCR [3]	1.0142	1.2643	0.6496	0.8297	1.2534	1.3499	0.7814	0.9689
DeepSoR [4]	0.9322	1.1563	0.6453	0.8239	1.0561	1.2034	0.7269	0.9008
SoMA	**0.8357**	**1.0943**	**0.6307**	**0.7957**	**1.0506**	**1.1890**	**0.7147**	**0.8838**

extending the proposed model with the aid of deep generative model (e.g., variational auto-encoder [5]) and further improving the recommendation performance will be interesting.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 61822601, 61773050, 61632004), Beijing Natural Science Foundation (Grant No. Z180006), National Key Research and Development Program (Grant No. 2017YFC1703506), and Fundamental Research Funds for the Central Universities (Grant No. 2019JBZ110).

Supporting information Appendixes A–C. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Andersen E B. Sufficiency and exponential families for discrete sample spaces. J Am Stat Assoc, 1970, 65: 1248–1255
2 Tang J L, Wang S, Hu X, et al. Recommendation with social dimensions. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence, 2016. 251–257
3 Wang X, He X N, Nie L Q, et al. Item silk road: recommending items from information domains to social users. In: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2017. 185–194
4 Fan W Q, Li Q, Cheng M. Deep modeling of social relations for recommendation. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2018. 8075–8076
5 Kingma D P, Welling M. Auto-encoding variational Bayes. 2013. ArXiv:1312.6114