Modeling Users' Exposure with Social Knowledge Influence and Consumption Influence for Recommendation

Jiawei Chen[1,2],Yan Feng[1,2],Martin Ester[3],Sheng Zhou[1],Chun Chen[1],Can Wang[1,2,*]
[1]College of Computer Science, Zhejiang University, China
[2]Zhejiang University-LianlianPay Joint Research Center
[3]School of Computing Science, Simon Fraser University, Canada
{sleepyhunt,fengyan}@zju.edu.cn;ester@cs.sfu.ca;{zhousheng_zju,chenc,wcan}@zju.edu.cn

ABSTRACT

Users' consumption behaviors are affected by both their personal preference and their *exposure* to items (i.e. whether a user knows the items). Most of the recent works in social recommendation assume that people share similar preference with their socially connected friends. However, this assumption may not hold due to the diversity of social relations, and modeling social influence on users' preference may not be suitable for *implicit feedback* data (i.e. whether a user has consumed certain items). Since users often share item information with their social relations, it will be less restrictive to model social influence on users' exposure to items. We notice that a user's exposure is affected by the exposure of the other users in his social communities and by the consumption of his connected friends. In this paper, we propose a novel social exposure-based recommendation model SoEXBMF by integrating two kinds of social influence on users' exposure, i.e. *social knowledge influence* and *social consumption influence*, into basic EXMF model for better recommendation performance. Furthermore, SoEXBMF uses Bernoulli distribution instead of Gaussian distribution in EXMF to better model the binary implicit feedback data. A variational inference method has been developed for the proposed SoEXBMF model to infer the posterior and make the recommendations. Extensive experiments on three real-world datasets demonstrate the superiority of our method over existing methods in various evaluation metrics.

CCS CONCEPTS

• **Information systems → Social recommendation**;

KEYWORDS

Social recommendation, Exposure, Implicit feedback, Graphic model

ACM Reference Format:
Jiawei Chen, Yan Feng, Martin Ester, Sheng Zhou, Chun Chen, Can Wang. 2018. Modeling Users' Exposure with Social Knowledge Influence, and Consumption Influence for Recommendation. In *The 27th ACM International Conference on Information and Knowledge Management (CIKM '18), October 22–26, 2018, Torino, Italy*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3269206.3271742

*Corresponding author: wcan@zju.edu.cn

CIKM '18, October 22–26, 2018, Torino, Italy

ACM ISBN 978-1-4503-6014-2/18/10 . . . $15.00
https://doi.org/10.1145/3269206.3271742

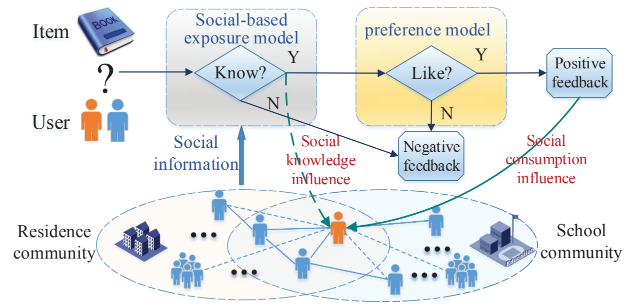


Figure 1: The generative process of implicit feedback with social influence on users' exposure

1 INTRODUCTION

With the exponential growth of information on e-commerce websites such as Amazon and Taobao and on-demand video streaming services such as Netflix and IMDB, recommender systems are drawing more attention from both academia and industry. Collaborative filtering (CF), as the prevalent recommendation model in these systems, infers user's preference and produces recommendations based on user's historical behaviors. However, traditional CF models are impeded by the data sparsity problem. That is, the number of items consumed by a user is often very small compared to the total number of items. The available historical feedback tends to be sparse (usually less than 0.1% [24]) and traditional CF models will suffer from severe performance degradation in such situations.

To overcome the limitation of traditional CF models, many social recommendation methods have been proposed by integrating social information into existing CF models for better performance. These methods mainly assume that people share similar preference with their socially connected friends [25, 29]. However, this assumption may not hold due to the diversity of social relations. For example, users in online social networks are linked for various reasons, e.g., alumni, colleagues, classmates or neighbors. Similar preference is not the exclusive motivation to get connected, and the connected users might keep diverse tastes. In fact, some researchers in social recommendation have noticed the gap between the social links and user preference similarity [2, 3], and thus proposed some sophisticated models to bridge this gap for better inferring user's preference in social networks. However, such sophisticated models may not be suitable for implicit feedback data, which contains only implicit information about user's consumption instead of explicit rating values. The one reason is that for implicit feedback data we do not have rating information to accurately analyze the effect of social networks on users' preference. The other reason is that there exist two different reasons for negative feedback: *unknown* or *dislike*. As users have only limited awareness of items, they may just not know items that they have not consumed. It is obviously more difficult to infer user's preference from implicit feedback. However, implicit feedback is more abundant than explicit feedback in practice [11] and the social recommendation methods that exploit implicit feedback will be more valuable.

To make better social recommendation for implicit feedback data, in this paper we consider using social information in a different way. Since users often share item information with their friends, it will be less restrictive to model social influence on user's exposure. This way, we can better interpret the reasons (unknown or dislike) for negative feedback and judge whether the user really dislike the item to improve recommendation performance. We explore the following two questions:

(1) Will a user's exposure be influenced by unconnected users? The answer is positive. Social networks enable effective information sharing. Thus item information will diffuse in a social network and a user's exposure may be affected by the knowledge (exposure) of the friends of friends. Furthermore, some recent works [21, 42] suggest that each of us belongs to some content-sharing communities. Thus the items mentioned by the users in our communities are more likely brought to our attention, even if they have no explicit connections with us. That is, a user's exposure is affected by the knowledge of both connected users and unconnected users in his communities. We name this influence as *social knowledge influence*, as denoted by dashed lines in Figure 1.

(2) Will a user's friends' consumption of an item increase the user's exposure to the item? Certainly! The items consumed by our friends are more likely to come to our attention than those that our friends were only exposed to. On the one hand, once a user consumes an item, it's more likely for him to share his personal comments on the item with his friends. On the other hand, e-commerce websites usually display the information about the consumption of user's friends, which boosts his exposure to these items. To emphasize this influence from one's friends' consumptions on his exposure, we name it as *social consumption influence*, as denoted by solid lines in Figure 1.

To summarize, we illustrate the generative process of implicit feedback for a specific user-item pair in Figure 1. The user develops his own exposure on the item with the influence from social networks, i.e. social knowledge influence and social consumption influence. After the user has learned the item, he will decide whether or not to consume the item based on his preference. Although both social knowledge influence and social consumption influence bear significant importance on user's exposure, no existing social recommendation methods take either of them into consideration. In this paper, we propose a novel social recommendation model named as SoEXBMF for implicit feedback data, which integrates both social knowledge influence and social consumption influence into the existing *exposure-based matrix factorization model* (EXMF) [15] to achieve better recommendation performance. Furthermore, to better model the binary implicit feedback data, we use Bernoulli distribution in SoEXBMF, which can better interpret the reasons (unknown or dislike) for negative feedback than the traditional Gaussian distribution in EXMF.

It is worthwhile to highlight the following contributions:

- We introduce the concepts of social knowledge influence and social consumption influence on user's exposure for improving social recommendation.
- We propose the generative probabilistic model SoEXBMF which integrates the two types of social influence on exposure into EXMF and models the binary implicit feedback with Bernoulli distribution instead of Gaussian distribution.
- We employ a Gaussian lower bound to deal with the unconjugated Bernoulli-logit structure and develop a variational inference method to infer the posterior for our SoEXBMF.
- Our experimental evaluation on three well-known benchmark datasets demonstrates that SoEXBMF consistently outperforms a range of state-of-the-art methods and analyzes the contributions of the different components of our method.

The rest of this paper is organized as follows. We briefly review related works in section 2. We give the problem definition and background in section 3. The exposure model with social influence is introduced in section 4. In section 5, we present the details of the SoEXBMF model. The experimental results and discussions are presented in section 6. Finally, we conclude the paper and present some directions for future work in section 7.

2 RELATED WORK

With the exponential growth of information generated on consumer review websites and e-commerce websites, recommender systems are drawing more attention from both academia and industry. Substantial works have been done about collaborative filtering (CF) model for its accuracy and scalability during the past two decades [10, 14, 23, 30–32]. Here, we review the most related works from two perspectives: one on the recommendation with implicit feedback and the other on the social recommendation.

Recommendation with implicit feedback. In implicit feedback settings, all the items, including the ones that a user did not consume, are taken into consideration. Negative feedback can be attributed to two reasons: *unknown*, or *dislike*. Thus to better infer user's preference, Weighted Matrix Factorization (WMF)[11], the standard factorization model for the implicit feedback data, selectively downweights the evidence of these negative feedback data. That is, WMF used a simple heuristic where all negative feedback data are equally downweighted vis-a-vis the positive feedback data. Also, some neural-based collaborative methods including CDAE[35], NCF[6] assign low confidence on the negative feedback by uniformly sampling a subset of the negative feedback to train their models. More recently, a new probabilistic model EXMF[15] incorporated user's exposure to items into the CF methods. In EXMF, user's exposure can be translated to the weight of the user-item interactions to downweight the negative feedback data automatically. As our method is based on EXMF, we will introduce EXMF model in details in section 3.

Some other pair-wise methods treat negative feedback data in a different way. Bayesian personalized ranking (BPR) [22] focuses on the learning of relative preferences and aims at maximizing the AUC objective function. They believe that user's preference on his consumed items is above un-consumed items. Similarly, the weighted approximated ranking pairwise (WARP) loss proposed in [34] optimize Precision@K. Also, there are some other works modeling implicit feedback with poisson distribution [3, 5] that can scale up to massive data sets.

Social Recommendation. Social information has been utilized to improve recommendation performance in recent works. These methods mainly assume that connected users will share similar preference [18, 38, 39]. Sorec [17], TrustMF [37], PSLF [25], jointly factorize rating matrix and trust (social) matrix by sharing a common latent user space. In [2, 3, 16, 26, 31, 36], users' feedback is considered as synthetic results of their preference and social influence. [13, 31] utilize a social regularization term to constrain user's latent preference close to his trusted friends. Also, some other researchers [33, 41] extend pair-wise BPR framework by further assuming that for all items with negative feedback, a user would prefer the items consumed by their friends over the rest.

Only one recent work SERec [29] in AAAI'18 explicitly employs social information to infer user's exposure to items. SERec extends EXMF to construct two social exposure-based CF models: (1) SERec-Re, in which the exposure matrix and the social relationship matrix are jointly factorized by sharing the common latent vectors of users; (2) SERec-Bo, in which a conditional model is used to heuristically infer user's exposure with homogeneous social influence from the exposure of his connected friends. However, although users' social relations are integrated into users' exposure in SERec, our method SoEXBMF differs from SERec in following aspects: (1) *social consumption influence* is explicitly modeled in SoEXBMF, while it is ignored by SERec in that they do not consider the boost of user's friends' consumption to his exposure. (2) *social knowledge influence* is explicitly modeled in SoEXBMF, while it is ignored by SERec in that they do not consider the influence between unconnected users in the common communities. (3) SoEXBMF explicitly considers heterogenous social influence while SERec treat social influence as homogeneous. (4) SoEXBMF uses Bernoulli distribution, which can better model the binary implicit feedback data than Guassian distribution in SERec. Overall, we summarize the comparison between the two methods in Table 2 and analyze the contributions of these four aspects in our experiments.

3 PROBLEM DEFINITION AND BACKGROUND

Suppose we have a recommender system with user set U (including n users) and item set I (including m items). The implicit feedback data is represented as $n \times m$ matrix X with entries x_{ij} denoting whether or not user i has consumed the item j. Social information indicates the connections between users. $\mathcal{T}_i$ denotes the set of connections (friends) of user i. The task of a recommender system can be stated as follows: recommending items for each user that are most likely to be consumed by him.

In the following we introduce EXMF [15], on which our proposed method is based. EXMF directly incorporates user's exposure to items into the CF methods. Firstly it generates the latent variable a_{ij}, which indicates whether user i has been exposed to item j. Then, it generates user's consumption x_{ij} based on a_{ij},

$$a_{ij} \sim Bernoulli(\eta_{ij}) \tag{1}$$

$$x_{ij}|a_{ij} = 1 \sim \mathcal{N}(u_i^\top v_j, \lambda_x^{-1}) \tag{2}$$

$$x_{ij}|a_{ij} = 0 \sim \delta_0 \tag{3}$$

where δ_0 denotes $p(x_{ij} = 0|a_{ij} = 0) = 1$; η_{ij} is the prior probability of exposure. When $a_{ij} = 0$, we have $x_{ij} = 0$, since the user does not know the item. When $a_{ij} = 1$, x_{ij} is modeled as the classic preference model and factorized by the latent vectors u_i and v_j, which respectively characterize the latent preferences of user i and latent attributes of item j. The expectation of user's exposure a_{ij} can be calculated as follows:

$$E[a_{ij}] = \frac{p(a_{ij} = 1)p(x_{ij}|a_{ij} = 1)}{\sum\limits_{a_{ij}=0,1} p(a_{ij})p(x_{ij}|a_{ij})} \tag{4}$$

That is, when $x_{ij} = 1$, $E[a_{ij}|x_{ij} = 1] = 1$, and when $x_{ij} = 0$,

$$E[a_{ij}|x_{ij} = 0] = \frac{\eta_{ij} \cdot \mathcal{N}(0|u_i^\top v_j, \lambda_x^{-1})}{(1 - \eta_{ij}) \cdot 1 + \eta_{ij} \cdot \mathcal{N}(0|u_i^\top v_j, \lambda_x^{-1})} \tag{5}$$

4 MODELING SOCIAL INFLUENCE ON USER'S EXPOSURE

In this section, we model two kinds of social influence on user's exposure: social knowledge influence and social consumption influence, and combine these two parts into an integrated exposure model.

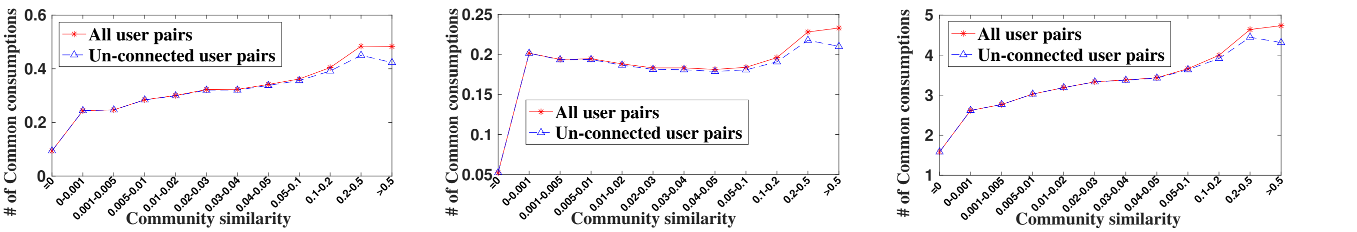
4.1 Social knowledge influence on exposure

Here we design an exposure model to capture the social influence between two users' exposure. As previously discussed, a user's exposure will be affected by the knowledge (exposure) of other users in his communities, including connected and unconnected users. To further validate the point, we conduct the following statistical analysis on three real-world social recommendation datasets: Ciao, Epinions, LastFM (Details about datasets are presented in section 6.1).

We employ the Mixed Membership stochastic blockmodel (MMSB)[1] [1] to discover communities in the social network and to learn a community distribution for each user. In Figure 2, we plot the average number of shared common consumptions vs. the (cosine) similarity of the community distribution, both for connected user pairs and for all user pairs. As shown in Figure 2, as users belong to more similar communities, they will share more common consumptions even if they are not directly connected. Also, we observe that the two curves in Figure 2 mostly overlap. The reason is that the number of un-connected user pairs is much larger than the connected user pairs. Thus, the influence between un-connected user pairs is also important in deriving user's exposure.

Motivated by the above analysis, we propose an exposure model based on users' community distribution to capture the social influence between two users' exposure. The proposed model generates a user's exposure to an item based on the average popularity of the item in all the communities the user belongs to. That is, if an item has been exposed to many members of a user's communities, he

[1] Svinet package: https://github.com/premgopalan/svinet. Note that the number of communities D can also be inferred by Svinet.



(a) On dataset Ciao (b) On dataset Epinions (c) On dataset LastFM

Figure 2: The average number of common consumptions between un-connected and all user pairs for varying community similarity.

will be more likely to know this item. Formally, a user's exposure has the following generative process:

$$a_{ij} \sim Bernoulli(\sigma(\theta_i^\top d_j)) \tag{6}$$

Where $\sigma(.)$ denotes the logistic function, θ_i the D-dimensional community distribution of user i, d_j the D-dimensional parameter of item j, which captures the popularity of item j in each community. In this paper, the community distribution is obtained using the Mixed-Membership stochastic blockmodel (MMSB). We remark that other community detection algorithms can also be used [4, 40]. However, the community detection algorithm is not the primary focus of this paper and is orthogonal to our proposed model.

To better understand the social influence between two user's exposure in this model, we formulate the following lemma. The proof is included in appendix A.1.

Lemma 1. *Given other users' exposure to the specific item j ($a_{\neg ij}$), the conditional probability of the exposure for user i is:*

$$p(a_{ij}|\theta, d_j, a_{\neg ij}) = Bernoulli\left(\sigma\left(\sum_{k\neq i} w^j_{k\to i}(a_{kj} - \frac{1}{2})\right)\right) \tag{7}$$

where $a_{\neg ij}$ denotes $a_{1j}, \ldots, a_{nj}$ but with a_{ij} omitted. $w^j_{k\to i} = \theta_i^\top M^{ij}\theta_k$ denotes the weight of social influence from user k on user i. M^{ij} is a positive definite matrix calculated by:

$$M^{ij} = (\sum_{k\neq i}(\frac{\sigma(\theta_k^\top d_j) - \frac{1}{2}}{\theta_k^\top d_j})\theta_k\theta_k^\top)^{-1} \tag{8}$$

According to lemma 1, the conditional probability of user's exposure depends on the weighted sum of the exposure of other users. That is, if users with strong impact on us know the specific item j, we are more likely exposed to it. The coefficients $w^j_{k\to i}$ indicate the strength of the influence from different users based on their community distribution. Further, $w^j_{k\to i}$ is bilinear form which can be expressed as the weighted average of the elements of M^{ij}: $w^j_{k\to i} = \sum_{1\leq p,q\leq D} \theta_{kp}\theta_{iq}M^{ij}_{pq}$. In practise, since users typically focus on a few communities, we can find the diagonal elements of M^{ij} are much larger than the other elements. For example, in the Ciao dataset, the minimum diagonal elements of M^{ij} is 0.02, while the maximum off-diagonal elements is 7.8e-5. thus, in our model the more similar community distributions between users correspond to a larger value of the influence strength, which results in more common items that the users consume, modeling the phenomenon in Figure 2.

4.2 Social consumption influence on exposure

The items consumed by our connected friends are more likely to come to our attention than those not consumed by them. Here we try to capture this phenomenon and propose an exposure model with the social influence from the friends' consumption. Formally, we model the social consumption influence as follows:

$$a_{ij} \sim Bernoulli(\sigma(\tau_{i0} + \sum_{k\in\mathcal{T}_i} \tau_{ik}x_{kj})) \tag{9}$$

where $\mathcal{T}_i$ denotes the set of friends of user i, and x_{kj} denotes the consumption of his friend user k. Also, the logistic function $\sigma(.)$ is employed to map the weighted sum to the probabilistic interval [0,1] for a stable and robust model. The user's exposure in Eq.(9) is modeled as a logistic regression based on the consumption of the user's friends. The regression parameter $\tau_{ik} > 0$ represents how much user i is influenced by the consumption of his friend user k. Intuitively, different social ties may have different influence strength (tie strength) [5, 28]. For example, the items consumed by our close friends are more likely to come to our attention than those from acquaintances. Our model captures this point and learns the personalized tie strength (τ_{ik}) from the data. Overall, as we can see from equation (9), the more and the closer friends (i.e., those with high value of τ_{ik}) have consumed the item, the greater the likelihood that the user will know the item.

4.3 Integrated social influence on exposure

Individuals will be affected by both the exposure of other users in their communities and by the consumption of friends. That is, on the one hand, each of us belongs to some content-sharing communities and the items mentioned in these communities are more likely brought to our attention. On the other hand, we are more likely to know the items that have been consumed by our friends. Thus both kinds of social influence are important for inferring users' exposure. We combine social consumption influence and social knowledge influence in an integrated social exposure model as follows in a Bernoulli-logistic structure:

$$a_{ij} \sim Bernoulli(\sigma(\theta_i^\top d_j + \tau_{i0} + \sum_{k\in\mathcal{T}_i} \tau_{ik}x_{kj})) \tag{10}$$

Eq.(10) indicates the larger the combined social influence, the more likely that the user will know the item.

5 SOCIAL EXPOSURE-BASED BINARY MATRIX FACTORIZATION

In this section, we will first introduce our improved exposure-based recommendation model EXBMF. Then, we will integrate the social exposure model from section 4 into EXBMF.

5.1 Exposure-based binary matrix factorization (EXBMF)

Considering implicit feedback data (x_{ij}) is a binary variable, it seems not a good choice to generate x_{ij} based on a continuous Gaussian distribution. Especially, for an exposure-based recommendation model, modeling the probability mass $p(x_{ij}|a_{ij} = 1)$ by a Gaussian distribution has the following drawback: as shown in Eq. (4) and (5), two conditional probabilities $p(x_{ij} = 0|a_{ij} = 1)$ and $p(x_{ij} = 0|a_{ij} = 0)$ need to be compared to infer user's exposure a_{ij}. If the probability mass $p(x_{ij} = 0|a_{ij} = 1)$ was replaced by the probability density function $N(0|u_i^\top v_j, \lambda_x^{-1})$, we will compare the mass function and density function. The value of a_{ij} is thus very sensitive to the selected Gaussian precision parameter λ_x and may not reflect the true exposure of users.

Thus, we replace the Gaussian distribution in the EXMF model by a Bernoulli distribution to generate x_{ij} and conduct matrix factorization on the Bernoulli parameters in our EXBMF model as follows:

$$x_{ij}|a_{ij} = 1 \sim Bernoulli(\sigma(z + u_i^\top v_j)) \tag{11}$$

$$x_{ij}|a_{ij} = 0 \sim \delta_0 \tag{12}$$

where δ_0 is a delta function with $p(x_{ij} = 0|a_{ij} = 0) = 1$, z is a global bias, and u_i is the latent K-dimensional vector of the preference of user i and v_j is the latent K-dimensional vector of the attribute of item j. Similar to EXMF, we employ simple matrix factorization to set off the effect of the Bernoulli model.

Exploring the reasons for negative feedback. With Bernoulli distribution on x_{ij}, our EXBMF can quantitatively explore the reasons for user's negative feedback (dislike, unknown or both). To facilitate the description, we introduce an auxiliary variable r_{ij}, which denotes whether user i likes the item j:

$$r_{ij} \sim Bernoulli(\sigma(z + u_i^\top v_j)) \tag{13}$$

$$x_{ij} = a_{ij} \cdot r_{ij} \tag{14}$$

Now, user's feedback x_{ij} can be considered as the product of two binary variables a_{ij} and r_{ij}. When $x_{ij} = 0$, we have three cases of the values of r_{ij} and a_{ij}: (1) $r_{ij} = 1, a_{ij} = 0$, (2) $r_{ij} = 0, a_{ij} = 1$, (3) $r_{ij} = 1, a_{ij} = 1$, which correspond to three cases of user's negative feedback: (1) The user likes the item but he does not know it. (2) The user knows the item but he dislikes it. (3) The user does not know the item and he dislikes it. In our EXBMF model, by inferring the joint posterior of a_{ij} and r_{ij}, we can quantitatively explore the reasons for user's negative feedback (dislike, unknown or both) and infer the probabilities for the three cases. These posterior probabilities can bring benefits in many applications. For example, these probabilities can help a company to find out the user-item pairs with high preference and low exposure to promote its products.

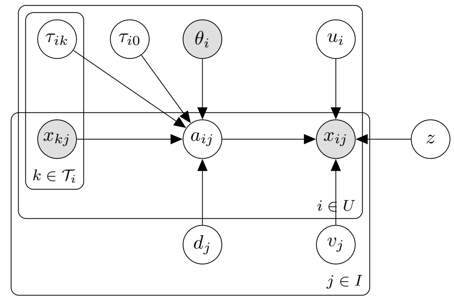


Figure 3: The graphical model of SoEXBMF.

5.2 Social exposure-based binary matrix factorization (SoEXBMF)

We now integrate the social exposure model into EXBMF to complete SoEXBMF. Here we introduce the zero-mean spheral Gaussian prior for penalizing the model complexity as proposed by recent works [15, 19, 29]. User's exposure (a_{ij}) to item is generated with social influence as defined in section 4. Implicit feedback data x_{ij} is generated based on user's exposure and preference according to our EXBMF model. Overall, the SoEXBMF model assumes the following generative process,corresponding to the graphical model presented in Figure 3:

(1) Draw the global bias: $z \sim \mathcal{N}(0, \lambda_z^{-1})$.
(2) For each user i:
  I. Draw bias: $\tau_{i0} \sim \mathcal{N}(0, \lambda_{\tau 0}^{-1})$.
  II. Draw the latent preference: $u_i \sim \mathcal{N}(0, \lambda_u^{-1}\mathbf{I})$.
  III. For each tie with friend k, draw the personalized tie strength: $\tau_{ik} \sim \mathcal{N}(0, \lambda_\tau^{-1})$.
(3) For each item j:
  I. Draw the item popularity in communities: $d_j \sim \mathcal{N}(0, \lambda_d^{-1}\mathbf{I})$.
  II. Draw the latent attribute: $v_j \sim \mathcal{N}(0, \lambda_v^{-1}\mathbf{I})$.
(4) For each user-item pair (i, j), draw users' exposure based on the synthetic influence of social network:
$a_{ij} \sim Bernoulli(\sigma(\theta_i^\top d_j + \tau_{i0} + \sum_{k \in \mathcal{T}_i} \tau_{ik} x_{kj}))$.
(5) For each user-item pair (i, j), draw the implicit feedback based on the exposure:
When the user know the item, draw the implicit feedback based on preference: $x_{ij}|a_{ij} = 1 \sim Bernoulli(\sigma(z + u_i^\top v_j))$.
When the user don't know the item, he will not consume it: $x_{ij}|a_{ij} = 0 \sim \delta_0$.

5.3 Variational inference

Since the exact posterior probability of SoEXBMF is not tractable to compute, we develop an efficient approximate method to compute the posterior based on variational inference [27]. The variational posterior distribution approximates the exact posterior of latent variables with a simpler, tractable distribution. The mean field theory motivates us to factorize these variational distributions into disjoint groups, and these variables are governed by their own variational parameters. Besides, we can specify the form of the factored variational distribution of each variable as same as its corresponding conditional [9]: the variational distributions of latent variables u, v, d, z, τ are Gaussian distributions, while a is a Bernoulli distribution. That is, we define variational distribution as follows:

$$
\begin{aligned}
q(u, v, z, d, \tau, a) = \mathcal{N}(z|\bar{z}, \tilde{z}) \prod_{j \in V} \Big(\mathcal{N}(v_j|\bar{v}_j, \tilde{v}_j) \mathcal{N}(d_j|\bar{d}_j, \tilde{d}_j) \Big) \\
\times \prod_{i \in U} \left(\mathcal{N}(u_i|\bar{u}_i, \tilde{u}_i) \mathcal{N}(\tau_{i0}|\bar{\tau}_{i0}, \tilde{\tau}_{i0}) \prod_{k \in \mathcal{T}_i} \mathcal{N}(\tau_{ik}|\bar{\tau}_{ik}, \tilde{\tau}_{ik}) \right) \\
\times \prod_{i \in U, j \in V} Bernoulli(a_{ij}|\bar{a}_{ij}) \qquad (15)
\end{aligned}
$$

where $\Phi = \{\{\bar{v}_j, \tilde{v}_j, \bar{d}_j, \tilde{d}_j\}_{j \in V}, \{\bar{u}_i, \tilde{u}_i, \bar{\tau}_{i0}, \tilde{\tau}_{i0}, \{\bar{\tau}_{ik}, \tilde{\tau}_{ik}\}_{k \in \mathcal{T}_i}\}_{i \in U}, \{\bar{a}_{ij}\}_{i \in U, j \in V}, \bar{z}, \tilde{z}\}$ are variational parameters that are adjusted so that $q(u, v, z, d, \tau, a)$ is as similar as possible to the posterior $p(u, v, z, d, \tau, a|X)$ by minimizing the KL divergence between them. Equivalently, variational parameter Φ can be optimized by maximizing the evidence of lower bound (ELBO) as follows:

$$
L(\Phi) = E_q[\log p(u, v, z, d, \tau, a, X)] - E_q[\log q(u, v, z, d, \tau, a) \qquad (16)
$$

Note that $L(\Phi)$ is analytically intractable since the Bernoulli-logit likelihood does not admit a conjugate prior. To address this problem we employ Gaussian lower bound on the logistic function [8, 12] as follow:

$$
\begin{aligned}
\sigma(x) \geqslant h(x, \xi) = \sigma(\xi) \exp((x - \xi)/2 - \lambda(\xi)(x^2 - \xi^2)) \\
\lambda(\xi) = \frac{1}{2\xi}(\sigma(\xi) - \frac{1}{2}) \qquad (17)
\end{aligned}
$$

which is a tight lower bound on the logistic function, with an additional parameter ξ. Thus, we can replace the logistic function of Bernoulli-logit likelihood for each x_{ij} and a_{ij} with an instantiation of (17) including additional parameters ξ^x_{ij} and ξ^a_{ij}. We obtain the following new lower bound:

$$
\begin{aligned}
\tilde{L}(\Phi, \xi) = \sum_{i \in U, j \in V} E_q \Big[h\left((2a_{ij} - 1)\alpha_{ij}, \xi^a_{ij}\right) + a_{ij} ln(\mathbb{I}[x_{ij} = 1]) \\
+ (1 - a_{ij}) h\left((2x_{ij} - 1)\beta_{ij}, \xi^x_{ij}\right) \Big] + E_q\left[\lambda_z z^2\right] \\
+ \sum_{i \in U} E_q \Big[\lambda_u u_i^2 + \lambda_{\tau 0} \tau_{i0}^2 + \lambda_\tau \sum_{k \in \mathcal{T}_i} \tau_{ik}^2 \Big] \\
+ \sum_{j \in V} E_q \Big[\lambda_v v_j^2 + \lambda_d d_j^2 \Big] - E_q\Big[q(u, v, z, d, \tau, a)\Big] \qquad (18)
\end{aligned}
$$

where $\alpha_{ij} = \theta_i^\top d_j + \tau_0 + \sum_{k \in \mathcal{T}_i} \tau_{ik} x_{kj}$ and $\beta_{ij} = u_i^\top v_j + z$ are the Bernoulli parameters of a_{ij} and x_{ij} respectively, and $\mathbb{I}[.]$ denotes the indicator function. We then use the coordinate ascent method to optimize the variational parameters iteratively by maximizing the lower bound $\tilde{L}(\Phi, \xi)$. In Algorithm 1, we present the pseudo-code of our inference algorithm, with details in appendix A.2.

Recommendations. Once the posterior is fit, recommendations are generated according to the probability that the user will consume the items as Eq.(41).

Algorithm 1 Variational inference of SoEXBMF

```
Initialize variational parameters randomly;
while not converge do
    for each user-item pair (i, j) do
        Update the parameters ξ^a_ij, ξ^x_ij and the variational parameters of
        a_ij. {Eq.(26)-(28)}
    end for
    Update the variational parameters of z. {Eq.(39)-(40)}
    for each user i do
        Update the variational parameters of u_i, τ_i0. {Eq.(29)-(32)}
        for each connection of the user i with the friend k do
            Update the variational parameters of τ_ik. {Eq.(33)-(34)}
        end for
    end for
    for each item j do
        Update the global variational parameters of v_j, d_j. {Eq.(35)-(38)}
    end for
end while
```

Table 1: Statistics of three datasets

Datasets	#Users	#Items	#Links	#User-item interactions
Ciao	7,375	11,166	111,781	147,814
Epinions	28,305	21,323	426,936	421,583
LastFM	1,892	2,828	25,434	71,426

6 EXPERIMENTS AND ANALYSES

In this section, we conduct experiments to evaluate the recommendation quality of SoEXBMF. Our experiments are intended to address the following questions:

(1) Is it beneficial to model user's exposure with the social consumption influence?
(2) Is it beneficial to model user's exposure with the social knowledge influence?
(3) Is it beneficial to model the heterogeneous social influence between users?
(4) Is it beneficial to model implicit feedback data with Bernoulli distribution instead of Gaussian distribution in exposure-based recommendation model?

6.1 Experimental protocol

Datasets. Three datasets Epinions[2], Ciao[3], LastFM[4] are used in our experiments. These datasets contain users' feedback and social relations. Specifically, the datasets Epinions and Ciao contain users' ratings on movies, while the dataset LastFM contains users' clicks on music. The dataset statistics are presented in Table 1. Similar to [7, 36], we preprocess the datasets so that all items have at least five interactions and "binarize" user's feedback into implicit feedback. That is, as long as there exists some user-item interactions (ratings or clicks), the corresponding implicit feedback is assigned a value of 1. Grid search and 5-fold cross validation are used to find the best parameters. That is, we set $K = 10$, $\lambda_u = \lambda_v = \lambda_z = \lambda_d = 1$, $\lambda_\tau = \lambda_{\tau 0} = 0.1$.

[2]http://www.trustlet.org/epinions
[3]http://www.cse.msu.edu/~tangjili/trust
[4]https://grouplens.org/datasets/hetrec-2011/

Table 2: The characteristics of the exposure-based compared methods.

Methods	EXMF	SERec-Bo	EXBMF	SERec-Bo-B	SoEXBMF-CH	SoEXBMF-C	SoEXBMF-K	SoEXBMF
Social influence	\	√	\	√	√	√	√	√
Consumption influence	\	\	\	\	√	√	\	√
Knowledge influence	\	\	\	\	\	\	√	√
Heterogenous influence	\	\	\	\	\	√	√	√
Bernoulli likelihood	\	\	√	√	√	√	√	√

Methods Compared. The following recommendation methods are tested.

- WMF [11]: The classic weighted matrix factorization model for implicit feedback data.
- SPF [3]: A social recommendation model that incorporates social influence with users' latent preference based on poisson factorization.
- EXMF [15]: A probabilistic model that directly incorporates user's exposure to items into traditional matrix factorization. EXMF does not utilize social information and choose an item-dependent prior of user's exposure.
- SERec-Bo[29]: A probabilistic model that extends the EXMF model with social influence on user's exposure. Note that in [29] the authors reported that the performance of SERec-Bo is consistently better than their other model SERec-Re. Thus, here we choose SERec-Bo as a comparison.
- EXBMF, SERec-Bo-B: To validate the superiority of employing Bernoulli likelihood rather than Gaussian likelihood for implicit feedback data, we design two comparisons EXBMF, SERec-Bo-B that modify EXMF, SERec-Bo in this way.
- SoEXBMF: Our complete social exposure-based recommendation model.
- SoEXBMF-K, SoEXBMF-C: To analyze the effect of each of the two components of social influence separately, we compare the methods SoEXBMF-K that just considers social knowledge influence on exposure and SoEXBMF-C that just considers social consumption influence on exposure.
- SoEXBMF-CH: To explore the benefit of modeling the heterogeneous social influence between users, we design another model SoEXBMF-CH, a special case of SoEXBMF-C where the different social ties are constrained to have the same tie strength.

We also present comparisons for these exposure-based methods in Table 2.

Evaluation Metrics. We adopt the following metrics to evaluate recommendation performance:

- Recall@K (Rec@K): This metric quantifies the fraction of consumed items that are in the top-K ranking list sorted by their estimated rankings. For each user i, we define $Rec(i)$ as the set of recommended items in top-K and $Con(i)$ as the set of consumed items in test data for user i. Then we have:

$$Recall@K = \frac{1}{|U|} \sum_{i \in U} \frac{|Rec(i) \cap Con(i)|}{|Con(i)|} \tag{19}$$

- Precision@K (Pre@K): This measures the fraction of the top-K items that are indeed consumed by the user:

$$Precision@K = \frac{1}{|U|} \sum_{i \in U} \frac{|Rec(i) \cap Con(i)|}{|Rec(i)|} \tag{20}$$

- Normalized Discounted Cumulative Gain (NDCG): This is widely used in information retrieval and it measures the quality of ranking through discounted importance based on positions. In recommendation, NDCG is computed as follow:

$$NDCG = \frac{1}{|U|} \sum_{i \in U} \frac{DCG_i}{IDCG_i} \tag{21}$$

where DCG_i is defined as follow and $IDCG_i$ is the ideal value of DCG_i coming from the best ranking.

$$DCG_i = \sum_{j \in Con(i)} \frac{1}{\log_2(rank_{ij} + 1)} \tag{22}$$

where $rank_{ij}$ represents the rank of the item j in the recommended list of the user i.

6.2 Results and Analyses

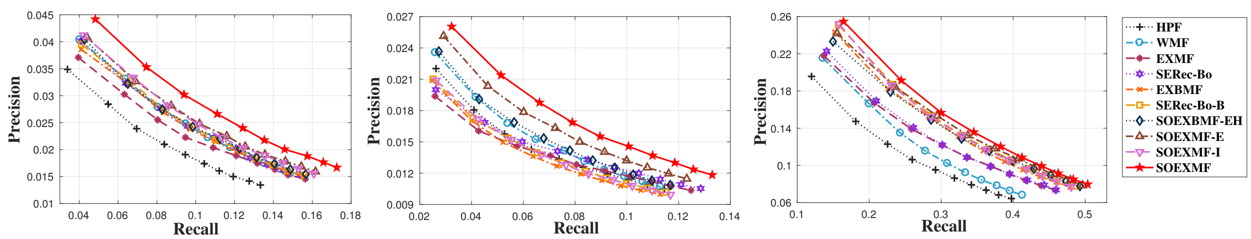
Table 3 presents the performance of the compared methods in terms of three evaluation metrics. The boldface font denotes the winner in that row. Overall, our proposed SoEXBMF model consistently outperforms all compared methods. Table 3 also shows the relative improvements achieved by SoEXBMF over the compared methods.

The effect of social influence. On the one hand, we observe that the methods with social knowledge influence outperform their corresponding special cases without it (SoEXBMF outperforms SoEXBMF-C, SoEXBMF-K outperforms EXBMF). The reason is that the information about items propagates along the social network. Users exposure to items is affected by the knowledge (exposure) of others including connected and unconnected users. This also can be seen from the performance gain of SOEXBMF-K versus SERec-Bo-B, which just models social influence between connected users. On the other hand, the methods (SoEXBMF, SoEXBMF-C, SoEXBMF-CH) that model social consumption influence on user's exposure achieve better performance than those that do not. This result validates our intuition that the items consumed by our friends are more likely to be exposed to us. Overall, we observe that SoEXBMF achieves best performance, which confirms the idea that combing both social consumption influence and social knowledge influence performs better than considering only one aspect.

The effect of Bernoulli likelihood. Note that the only difference between EXBMF and EXMF is that EXBMF employs Bernoulli likelihood for implicit feedback data while EXMF employs Gaussian likelihood. We can observe that EXBMF, with very few exceptions, outperforms EXMF. Especially in LastFM, the improvement is apparent: 11.7% in terms of precision, 11.1% in terms of recall and 4.5% in terms of NDCG. These results demonstrate the effectiveness of Bernoulli likelihood for the implicit feedback data. Similar results can be seen from the performance of SERec-Bo-B vs. SERec-Bo.

Table 3: The performance metrics of the compared algorithms. The boldface font denotes the winner in that row, and the second number in a cell shows the relative performance gain of SoEXBMF compared to that method.

Methods		SPF	WMF	EXMF	SERec-Bo	EXBMF	SERec-Bo-B	SoEXBMF-CH	SoEXBMF-C	SoEXBMF-K	SoEXBMF
Ciao	Pre@5	0.0349	0.0403	0.0371	0.0402	0.0386	0.0395	0.0405	0.0408	0.0412	**0.0442**
	Impv	26.6%	9.6%	19.1%	10.0%	14.5%	11.8%	9.2%	8.4%	7.2%	-
	Rec@5	0.0339	0.0401	0.0395	0.0409	0.0410	0.0411	0.0427	0.0441	0.0423	**0.0483**
	Impv	42.3%	20.4%	22.1%	18.2%	17.6%	17.4%	13.0%	9.5%	14.2%	-
	NDCG	0.2004	0.2133	0.2139	0.2145	0.2149	0.2159	0.2152	0.2168	0.2170	**0.2239**
	Impv	11.7%	4.9%	4.6%	4.4%	4.2%	3.7%	4.0%	3.3%	3.2%	-
Epinions	Pre@5	0.0220	0.0236	0.0194	0.0200	0.0209	0.0210	0.0237	0.0251	0.0209	**0.0260**
	Impv	18.3%	10.6%	34.4%	30.3%	24.9%	24.0%	10.1%	3.6%	24.8%	-
	Rec@5	0.0263	0.0260	0.0259	0.0262	0.0248	0.0251	0.0274	0.0292	0.0263	**0.0323**
	Impv	23.0%	24.4%	24.8%	23.5%	30.2%	28.7%	18.1%	10.7%	22.7%	-
	NDCG	0.1848	0.1819	0.1868	0.1884	0.1840	0.1848	0.1854	0.1885	0.1851	**0.1944**
	Impv	5.2%	6.9%	4.1%	3.2%	5.7%	5.2%	4.9%	3.1%	5.0%	-
LastFM	Pre@5	0.1957	0.2152	0.2180	0.2228	0.2435	0.2509	0.2333	0.2420	0.2520	**0.2547**
	Impv	30.2%	18.4%	16.8%	14.3%	4.6%	1.5%	9.2%	5.3%	1.1%	-
	Rec@5	0.1199	0.1358	0.1369	0.1409	0.1521	0.1575	0.1497	0.1554	0.1583	**0.1644**
	Impv	37.1%	21.1%	20.1%	16.7%	8.1%	4.4%	9.8%	5.8%	3.9%	-
	NDCG	0.4397	0.4601	0.4695	0.4707	0.4905	0.4959	0.4889	0.4935	0.4944	**0.5033**
	Impv	14.5%	9.4%	7.2%	6.9%	2.6%	1.5%	2.9%	2.0%	1.8%	-



(a) On dataset Ciao (b) On dataset Epinions (c) On dataset LastFM

Figure 4: Precision@K vs. Recall@K, where K ranges from 5 to 50

Heterogenous social influence. As we can see from Table 3, SoEXBMF-C with personalized tie strength consistently achieves better performance than its special case SoEXBMF-CH with the same tie strength. These results support our intuition that different ties may have different influence strength. For example, the items consumed by our close friends that have high frequency of interactions, are more likely to come to our attention than those items consumed by our acquaintances.

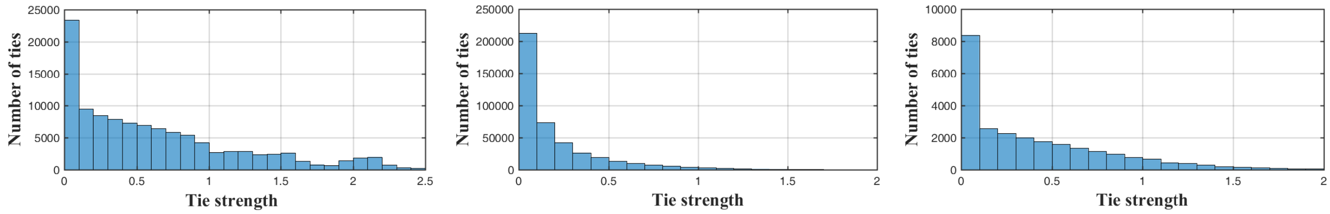
To further interpret the personalized tie strength in our model, we present the distribution of tie strength $\bar{\tau}_{ik}$ in Figure 5. These three plots are very consistent with those plots presented in [20], a research exploring tie strength in mobile communication networks. Figure 5 shows that tie strength has a fairly diverse distribution in all three datasets. Although there exist a large number of ties with weak strength ($t_{ik} < 0.1$), the numbers of ties with much stronger strength varying from 0.5 to 2.5 are nonnegligible. This confirms that different ties may indeed have different tie strength. Thus, explicitly modeling personalized tie strength in social recommendation is necessary.

Recall and Precision. Figure 4 depicts Recall (X-axis) vs. Precision (Y -axis) of the ten compared recommendation methods. Data points from left to right on each line were calculated at different values of K, ranging from 5 to 50. Clearly, the closer the line is to the top right corner (of the plot area), the better the algorithm is: which indicates that both recall and precision are high. We observe that SoEXBMF again clearly outperforms other methods, consistent with the findings in Table 3.

7 CONCLUSIONS

In this paper, we present a novel social recommendation method SoEXBMF for implicit feedback data, which models the social knowledge influence and social consumption influence on users' exposure. For the knowledge influence, we consider that users' exposure is affected by the knowledge (exposure) of the connected and unconnected users in their communities. For the consumption influence, we consider the heterogeneous social influence from connected users' consumption and the personalized tie strength can be learned in our model. Besides, in our model we utilize Bernoulli distribution instead of traditional Gaussian distribution to generate implicit feedback data, which can better interpret the reasons (unknown or dislike) of negative feedback. The experimental results on three real-world datasets show that SoEXBMF consistently outperforms existing methods in various metrics.

In the future, we plan to explore venue recommendation. In venue recommendation, a user's negative feedback can be attributed to three reasons: (1) he does not know it; (2) he dislikes it; (3) the



(a) On dataset Ciao (b) On dataset Epinions (c) On dataset LastFM

Figure 5: Histograms of tie strength $\bar{\tau}_{ik}$

venue is too far to visit. We can extend our SoEXBMF to model these three aspects for better recommendation performance.

ACKNOWLEDGMENTS

This work is supported by Zhejiang University-LianlianPay Joint Research Center and Zhejiang Provincial Key Research and Development Plan (Grant no. 2017C01012).

A APPENDIX

A.1 Proof of lemma 1

Formally, given other users exposure on the specific item j ($a_{\neg ij}$: $a_{1j}, \ldots, a_{nj}$ but with a_{ij} omitted), the optimal value of d_j need to maximize the following likelihood:

$$L = \sum_{k \neq i} \log Bernoulli(a_{kj} | \sigma(\theta_k^\top d_j)) \tag{23}$$

Note that L is convex on the variable d_j and thus $\frac{\partial L}{\partial d_j} = 0$, we have:

$$\sum_{k \neq i} \theta_k (a_{kj} - \frac{1}{2}) = \sum_{k \neq i} \theta_k (\sigma(\theta_k^\top d_j) - \frac{1}{2}) = \sum_{k \neq i} (\frac{\sigma(\theta_k^\top d_j) - \frac{1}{2}}{\theta_k^\top d_j}) \theta_k \theta_k^\top d_j$$

That is,

$$d_j = M^{ij} \sum_{k \neq i} \theta_k (a_{kj} - \frac{1}{2})$$

$$M^{ij} = (\sum_{k \neq i} (\frac{\sigma(\theta_k^\top d_j) - \frac{1}{2}}{\theta_k^\top d_j}) \theta_k \theta_k^\top)^{-1} = (\Theta_{\neg i} \Lambda^{ij} \Theta_{\neg i}^\top)^{-1} \tag{24}$$

Where $\Theta_{\neg i} = [\theta_1, \theta_2, ..\theta_{k \neq i}]$, Λ^{ij} denotes diagonal matrix with diagonal elements $\Lambda_{kk}^{ij} = \frac{\sigma(\theta_k^\top d_j) - \frac{1}{2}}{\theta_k^\top d_j}$. Here we note that the number of users n is much larger than the number of communities D, and thus believe that the matrix $\Theta_{\neg i}$ is full rank. If not, we can reduce the dimension to release the coupling. That is, for arbitrary non-zero D-dimensional vector v, we have $\Theta_{\neg i} v \neq 0$. Thus, we can find the matrix $\Theta_{\neg i} \Lambda^{ij} \Theta_{\neg i}^\top$ is a positive definite matrix, since the diagonal elements of Λ^{ij} are positive and the inequality $v^\top \Theta_{\neg i} \Lambda^{ij} \Theta_{\neg i}^\top v > 0$ holds. Thus, the inverse matrix M^{ij} is positive definite too.

Then, we have the following conditional distribution of user's exposure based on equations(6) and (24):

$$p(a_{ij} | \theta, a_{\neg ij}) = Bernoulli\left(\sigma\left(\sum_{k \neq i} (\theta_i^\top M^{ij} \theta_k)(a_{kj} - \frac{1}{2})\right)\right) \tag{25}$$

A.2 Variational Inference of SoEXBMF

By calculating the gradient of $\tilde{L}(\Phi, \xi)$ on each variational parameter, we can get the optimal closed solution based on other variational parameters. Thus, in each step, we can update these variational parameters as follows:

$$\xi_{ij}^x = \sqrt{Tr((\tilde{u}_i + \bar{u}_i \bar{u}_i^\top)(\tilde{v}_j + \bar{v}_j \bar{v}_j^\top)) + \tilde{z} + \bar{z}^2 + 2\bar{z}\bar{u}_i^\top \bar{v}_j} \tag{26}$$

$$\xi_{ij}^a = \sqrt{\bar{\alpha}_i j^2 + \theta_i^\top \tilde{d}_j \theta_i + \tilde{\tau}_{i0} + \sum_{k \in \mathcal{T}_i} \tilde{\tau}_{ik} x_{kj}} \tag{27}$$

$$\bar{a}_{ij} = \begin{cases} 1, x_{ij} = 1 \\ \frac{\exp(\rho_+(\xi_{ij}^a, \bar{\alpha}_{ij}) + \rho_-(\xi_{ij}^x, \bar{\beta}_{ij}))}{\exp(\rho_+(\xi_{ij}^a, \bar{\alpha}_{ij}) + \rho_-(\xi_{ij}^x, \bar{\beta}_{ij})) + \exp(\rho_-(\xi_{ij}^a, \bar{\alpha}_{ij}))}, x_{ij} = 0 \end{cases} \tag{28}$$

$$\tilde{u}_i = (\lambda_u I + \sum_{j \in V} 2\bar{a}_{ij} \lambda_{ij}^x (\bar{v}_j \bar{v}_j^\top + \tilde{v}_j))^{-1} \tag{29}$$

$$\bar{u}_i = \tilde{u}_i \sum_{j \in V} \bar{a}_{ij} \bar{v}_j (\frac{2x_{ij} - 1}{2} - 2\lambda_{ij}^x \bar{z}) \tag{30}$$

$$\tilde{\tau}_{i0} = (\lambda_{\tau 0} I + \sum_{j \in V} 2\lambda_{ij}^a)^{-1} \tag{31}$$

$$\bar{\tau}_{i0} = \tilde{\tau}_{i0} \sum_{j \in V} (\frac{2a_{ij} - 1}{2} - 2\lambda_{ij}^a (\theta_i^\top d_j + \sum_{k \in \mathcal{T}_i} \bar{\tau}_{ik} x_{kj})) \tag{32}$$

$$\tilde{\tau}_{ik} = (\lambda_\tau I + \sum_{j \in V} 2\lambda_{ij}^a x_{kj})^{-1} \tag{33}$$

$$\bar{\tau}_{ik} = \tilde{\tau}_{ik} \sum_{j \in V} x_{kj} (\frac{2a_{ij} - 1}{2} - 2\lambda_{ij}^a (\theta_i^\top d_j + \bar{\tau}_{i0} + \sum_{l \in \mathcal{T}_i, l \neq k} \bar{\tau}_{il} x_{lj})) \tag{34}$$

$$\tilde{v}_j = (\lambda_v I + \sum_{i \in U} 2\bar{a}_{ij} \lambda_{ij}^x (\bar{u}_i \bar{u}_i^\top + \tilde{u}_i))^{-1} \tag{35}$$

$$\bar{v}_j = \tilde{v}_j \sum_{i \in U} \bar{a}_{ij} \bar{u}_i (\frac{2x_{ij} - 1}{2} - 2\lambda_{ij}^x \bar{z}) \tag{36}$$

$$\tilde{d}_j = (\lambda_d I + \sum_{i \in U} 2\lambda_{ij}^a \theta_i \theta_i^\top)^{-1} \tag{37}$$

$$\bar{d}_j = \tilde{d}_j \sum_{i \in U} \theta_i (\frac{2a_{ij} - 1}{2} - 2\lambda_{ij}^a (\bar{\tau}_{i0} + \sum_{k \in \mathcal{T}_i} \bar{\tau}_{ik} x_{kj})) \tag{38}$$

$$\tilde{z} = (\lambda_z I + \sum_{j \in V, i \in U} 2\bar{a}_{ij} \lambda_{ij}^x)^{-1} \tag{39}$$

$$\bar{z} = \tilde{z} \sum_{j \in V, i \in U} \bar{a}_{ij} (\frac{2x_{ij} - 1}{2} - 2\lambda_{ij}^{x} \bar{u}_i^{\top} \bar{v}_j) \tag{40}$$

Where $Tr(.)$ denotes the trace of the matrix, $\rho_{+}(a,b) = log\sigma(a) + 0.5 \times (b-a)$ and $\rho_{-}(a,b) = log\sigma(a) + 0.5 \times (-b-a)$, where $\bar{\alpha}_{ij} = \theta_i^{\top} \bar{d}_j + \bar{\tau}_{i0} + \sum_{k \in \mathcal{T}_i} \bar{\tau}_{ik} x_{kj}$ and $\bar{\beta}_{ij} = \bar{u}_i^{\top} \bar{v}_j + \bar{z}$ are the expectation of the Bernoulli parameters of a_{ij} and x_{ij}.

Once the posterior is fit, recommendations generated according to the probability that the user will consume the items as follows:

$$\begin{aligned} E_q[x_{ij}] &= E_q[p(x_{ij} = 1)] = E_q[p(x_{ij} = 1 | a_{ij} = 1) p(a_{ij} = 1)] \\ &= \exp(\rho_{+}(\xi_{ij}^{a}, \bar{\alpha}_{ij}) + \rho_{+}(\xi_{ij}^{x}, \bar{\beta}_{ij})) \end{aligned} \tag{41}$$

REFERENCES

[1] Edoardo M Airoldi, David M Blei, Stephen E Fienberg, and Eric P Xing. 2008. Mixed membership stochastic blockmodels. *Journal of Machine Learning Research* 9, Sep (2008), 1981–2014.

[2] Y Bao, H Fang, and J Zhang. 2014. Leveraging decomposed trust in probabilistic matrix factorization for effective recommendation. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*. 350.

[3] Allison JB Chaney, David M Blei, and Tina Eliassi-Rad. 2015. A probabilistic model for using social networks in personalized item recommendation. In *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 43–50.

[4] Nan Du, Bin Wu, Xin Pei, Bai Wang, and Liutong Xu. 2007. Community detection in large-scale social networks. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*. ACM, 16–25.

[5] Prem Gopalan, Jake M. Hofman, and David M. Blei. 2015. Scalable Recommendation with Hierarchical Poisson Factorization. In *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence, UAI 2015, July 12-16, 2015, Amsterdam, The Netherlands*. 326–335.

[6] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 173–182.

[7] Xiangnan He, Hanwang Zhang, Min-Yen Kan, and Tat-Seng Chua. 2016. Fast matrix factorization for online recommendation with implicit feedback. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM, 549–558.

[8] José Miguel Hernández-Lobato, Neil Houlsby, and Zoubin Ghahramani. 2014. Stochastic inference for scalable probabilistic modeling of binary matrices. In *International Conference on Machine Learning*. 379–387.

[9] Matthew D Hoffman, David M Blei, Chong Wang, and John Paisley. 2013. Stochastic variational inference. *The Journal of Machine Learning Research* 14, 1 (2013), 1303–1347.

[10] Chin Chi Hsu, Mi Yen Yeh, and Shou de Lin. 2018. A General Framework for Implicit and Explicit Social Recommendation. *IEEE Transactions on Knowledge and Data Engineering* 14, 8 (2018). https://doi.org/10.1109/TKDE.2018.2821174

[11] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*. Ieee, 263–272.

[12] Tommi S Jaakkola and Michael I Jordan. 2000. Bayesian parameter estimation via variational methods. *Statistics and Computing* 10, 1 (2000), 25–37.

[13] Mohsen Jamali and Martin Ester. 2010. A matrix factorization technique with trust propagation for recommendation in social networks. In *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 135–142.

[14] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42, 8 (2009).

[15] Dawen Liang, Laurent Charlin, James McInerney, and David M Blei. 2016. Modeling user exposure in recommendation. In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 951–961.

[16] Hao Ma, Irwin King, and Michael R Lyu. 2009. Learning to recommend with social trust ensemble. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 203–210.

[17] Hao Ma, Haixuan Yang, Michael R Lyu, and Irwin King. 2008. Sorec: social recommendation using probabilistic matrix factorization. In *Proceedings of the 17th ACM conference on Information and knowledge management*. ACM, 931–940.

[18] Hao Ma, Dengyong Zhou, Chao Liu, Michael R Lyu, and Irwin King. 2011. Recommender systems with social regularization. In *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 287–296.

[19] Andriy Mnih and Ruslan R Salakhutdinov. 2008. Probabilistic matrix factorization. In *Advances in neural information processing systems*. 1257–1264.

[20] J-P Onnela, Jari Saramäki, Jorkki Hyvönen, György Szabó, David Lazer, Kimmo Kaski, János Kertész, and A-L Barabási. 2007. Structure and tie strengths in mobile communication networks. *Proceedings of the national academy of sciences* 104, 18 (2007), 7332–7336.

[21] Gergely Palla, Imre Derényi, Illés Farkas, and Tamás Vicsek. 2005. Uncovering the overlapping community structure of complex networks in nature and society. *Nature* 435, 7043 (2005), 814–818.

[22] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*. AUAI Press, 452–461.

[23] Ruslan Salakhutdinov and Andriy Mnih. 2007. Probabilistic Matrix Factorization.. In *Nips*, Vol. 1. 2–1.

[24] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*. ACM, 285–295.

[25] Yelong Shen and Ruoming Jin. 2012. Learning personal+ social latent factor model for social recommendation. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1303–1311.

[26] Jiliang Tang, Huiji Gao, and Huan Liu. 2012. mTrust: discerning multi-faceted trust in a connected world. In *Proceedings of the fifth ACM international conference on Web search and data mining*. ACM, 93–102.

[27] Martin J Wainwright and Michael I Jordan. 2008. Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning* 1, 1-2 (2008), 1–305.

[28] Menghan Wang, Xiaolin Zheng, Yang Yang, and Kun Zhang. 2017. Collaborative Filtering with Social Exposure: A Modular Approach to Social Recommendation. *arXiv preprint arXiv:1711.11458* (2017).

[29] Menghan Wang, Xiaolin Zheng, Yang Yang, and Kun Zhang. 2018. Collaborative Filtering With Social Exposure: A Modular Approach to Social Recommendation. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, February 2-7, 2018*.

[30] Xin Wang, Roger Donaldson, Christopher Nell, Peter Gorniak, Martin Ester, and Jiajun Bu. 2016. Recommending Groups to Users Using User-Group Engagement and Time-Dependent Matrix Factorization. *Proceedings of the 30th Conference on Artificial Intelligence (AAAI 2016)* (2016), 1331–1337.

[31] Xin Wang, Steven CH Hoi, Martin Ester, Jiajun Bu, and Chun Chen. 2017. Learning personalized preference of strong and weak ties for social recommendation. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1601–1610.

[32] Xin Wang, Steven CH Hoi, Chenghao Liu, and Martin Ester. 2017. Interactive social recommendation. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. ACM, 357–366.

[33] Xin Wang, Wei Lu, Martin Ester, Can Wang, and Chun Chen. 2016. Social Recommendation with Strong and Weak Ties. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, 5–14.

[34] Jason Weston, Samy Bengio, and Nicolas Usunier. 2011. Wsabie: Scaling up to large vocabulary image annotation. In *IJCAI*, Vol. 11. 2764–2770.

[35] Yao Wu, Christopher DuBois, Alice X Zheng, and Martin Ester. 2016. Collaborative denoising auto-encoders for top-n recommender systems. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*. ACM, 153–162.

[36] Lin Xiao, Zhang Min, Zhang Yongfeng, Liu Yiqun, and Ma Shaoping. 2017. Learning and transferring social and item visibilities for personalized recommendation. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. ACM, 337–346.

[37] Bo Yang, Yu Lei, Dayou Liu, and Jiming Liu. 2013. Social collaborative filtering by trust. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*. AAAI Press, 2747–2753.

[38] Xiwang Yang, Harald Steck, and Yong Liu. 2012. Circle-based recommendation in online social networks. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1267–1275.

[39] Weilong Yao, Jing He, Guangyan Huang, and Yanchun Zhang. 2014. Modeling dual role preferences for trust-aware recommendation. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*. ACM, 975–978.

[40] Zhen Zhang, Hongxia Yang, Jiajun Bu, Sheng Zhou, Pinggang Yu, Jianwei Zhang, Martin Ester, and Can Wang. 2018. ANRL: Attributed Network Representation Learning via Deep Neural Networks.. In *IJCAI*. 3155–3161.

[41] Tong Zhao, Julian McAuley, and Irwin King. 2014. Leveraging social connections to improve personalized ranking for collaborative filtering. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. ACM, 261–270.

[42] Tao Zhou. 2011. Understanding online community user participation: a social influence perspective. *Internet research* 21, 1 (2011), 67–81.