

Understanding Users' Budgets for Recommendation with Hierarchical Poisson Factorization

Yunhui Guo, Congfu Xu *, Hanzhang Song and Xin Wang

Institute of Artificial Intelligence, College of Computer Science, Zhejiang University, China
 {gyhui, xucongfu, hz_song}@zju.edu.cn, cswangxinm@gmail.com

Abstract

People consume and rate products in online shopping websites. The historical purchases of customers reflect their personal consumption habits and indicate their future shopping behaviors. Traditional preference-based recommender systems try to provide recommendations by analyzing users' feedback such as ratings and clicks. But unfortunately, most of the existing recommendation algorithms ignore the budget of the users. So they cannot avoid recommending users with products that will exceed their budgets. And they also cannot understand how the users will assign their budgets to different products. In this paper, we develop a generative model named collaborative budget-aware Poisson factorization (CBPF) to connect users' ratings and budgets. The CBPF model is intuitive and highly interpretable. We compare the proposed model with several state-of-the-art budget-unaware recommendation methods on several real-world datasets. The results show the advantage of uncovering users' budgets for recommendation.

1 Introduction

Online market is expanding to an unparalleled scale. Facilitated by websites such as Amazon.com and Alibaba.com, people more and more rely on online shopping and consume millions of products every day. This is accompanied with the rapid increase of the online retail. In 2010, the Internet economy accounts for 4.7% of the US gross domestic product (GDP) while in 2016 the contribution has increased to 5.4%¹. And the proportion of the online retail tends to continue to rise in the years ahead. This phenomenon inevitably has a big impact on every aspects of our life.

With the growth of the online market, there is an urgent need for intelligent recommendation of products for customers. Most of the online shopping websites leverage recommender systems to record users' usage data such as ratings and clicks to uncover their latent preferences. However, most of the recommendation algorithms are budget-unaware, in other words, they cannot avoid recommending

users with products that exceed their budgets and they also cannot understand how users will assign their budgets to different products. According to [Du and Kamakura, 2008], consumers allocate their budgets by considering both the “*whether to spend*” and the “*how much to spend*” decisions to maximize the utility function. Thus the consumers' decisions imply an unified preference structure for different products within a category. In this paper, we propose a model called collaborative budget-aware Poisson factorization (CBPF) to consider users' ratings and budgets simultaneously. With the understanding of how users' budgets affect their purchasing behaviors, we can provide more intelligent recommendations. Few works in recommender systems community consider the monetary limit of users. To the best of our knowledge, this is the first work that considers modeling users' budgets in a generative way for product recommendation.

Why try to model users' budgets? Unlike the category of books or movies, price plays a more important role in affecting users' purchases of commercial goods such as cellphones or jewelries. As pointed out by [Du and Kamakura, 2008; Kooti *et al.*, 2016], users usually have finite budgets and they will assign more budgets to the products that they prefer. So to understand how the users will divide their budgets across different products is of great practical importance. Moreover, some users may put more budgets on certain kinds of product attributes. For example, some customers may willing to pay extra budgets for a larger cellphone screen compared with a powerful battery. All these phenomena demonstrate the rationality and importance of uncovering users' budgets for intelligent item recommendation. This is a blind spot of traditional recommender systems.

Why model with Poisson factorization? Hierarchical Poisson factorization (HPF) [Gopalan *et al.*, 2013] is an effective item recommendation model for users' ratings. It has been used for content-based recommendation [Gopalan *et al.*, 2014], social-based recommendation [Chaney *et al.*, 2015] and dynamic recommendation [Charlin *et al.*, 2015]. There are three reasons for modeling users' budgets based on HPF. First is that HPF can be interpreted as a two-stage process that the user first decides a budget, then allocates this budget across different products. This interpretation makes HPF intuitive to model the process of users' allocation of budgets. Second is that previous works [Gopalan *et al.*, 2013; Chaney

* Corresponding author.

¹<https://www.bcg.com/documents/file100409.pdf>

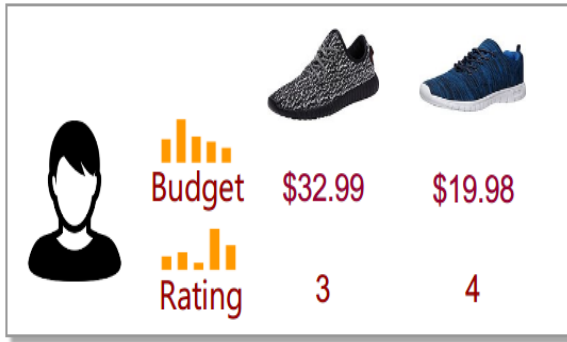


Figure 1: A simple illustration of users' budget preference and rating preference. The user's budget preference corresponds to the money the user spent on the product and the user's rating preference corresponds to the rating the user gave to the product. The orange bars represent the latent vector representations of users' preferences.

et al., 2015] point out HPF is better than Gaussian matrix factorization based models [Salakhutdinov and Mnih, 2008a; 2008b] in estimating users' activity levels. Third is that the inference of HPF is generally fast due to it only relies on the observed cells of the feedback matrix.

Case study: In Figure 1, we take a user's purchasing records as an example. The user bought two shoes from Amazon.com. He spent more on the first one but rated it lower. So from the perspective of preference-based recommender systems, the user preferred the second shoe to the first. On the other hand, from the perspective of budget-aware recommendation, since the user spent more on the first shoe, he preferred the first shoe to the second. That means the first shoe has some properties that the user preferred and decided to allocate more funds. This reflects that the products the users actually bought do not necessarily cost less. If the user spent a lot on a product, this means the user preferred the product to other products. So in general users will buy products that they both prefer and afford. The ratings the user given to the products and the budget the user allocated to the products are both indicators of the user's inclinations. Thus for intelligent recommendation, it is crucial to take both perspectives into consideration. The orange bars represent the latent vector representations of users' preferences. The height of the bars means what attribute the user will assign more budget and how the user will rate the attribute accordingly.

Summary and contributions: In this paper, we develop a generative model named collaborative budget-aware Poisson factorization (CBPF) to connect users' ratings and budgets simultaneously. The model is intuitive and highly interpretable. The contributions of the paper are summarized below:

- We propose a generative model to connect users' ratings and budgets in a cohesive manner.
- We develop an efficient variational inference optimization algorithm to approximate the posterior distribution of the model.
- The experimental results reveal the advantage of the proposed model compared with state-of-the-art budget-unaware recommendation models.

2 Related Work

Recommender systems have drawn more attentions from the machine learning community since the emergence of the e-commerce [Resnick and Varian, 1997]. Various approaches have been proposed to provide better product recommendations. Among them, collaborative filtering [Goldberg *et al.*, 1992; Koren, 2008; Koren and Bell, 2011] is a leading technique which tries to recommend a user with products by analysing similar users' records. Gaussian matrix factorization models [Salakhutdinov and Mnih, 2008a; 2008b] were proposed recently to handle extremely large datasets and to deal with cold-start users who have very few ratings. In matrix factorization models, users and items are represented as low-dimensional vectors. The latent user vectors indicate user preferences and latent item vectors indicate item attributes. This representation is also called latent factor model [Koren, 2008; Koren *et al.*, 2009].

In order to better handle the data sparsity problem and the long-tail of users and items, hierarchical Poisson factorization [Gopalan *et al.*, 2013] was proposed recently to be an alternative to Gaussian matrix factorization models. The main problem with Gaussian matrix factorization models is that they systematically overestimate users' activity levels [Gopalan *et al.*, 2013]. Recently, many variants of hierarchical Poisson factorization model [Gopalan *et al.*, 2014; Chaney *et al.*, 2015; Charlin *et al.*, 2015] were developed to provide more flexible item recommendations.

Understanding how budgets affect users' purchasing decisions is an important topic in the field of economics [Kao *et al.*, 2001; Gourinchas and Parker, 2002; Du and Kamakura, 2008]. However, all these works are taken from an economical perspective. Based on our analysis above, uncovering users' budgets can also be an important issue in the field of recommender systems. However, few works consider incorporating budgets into recommender systems. [Xie *et al.*, 2010] considers the situation that users choose a travel package on a limited budget. To the best of our knowledge, this paper is the first one that works on incorporating users' budgets into recommender systems in a generative way.

3 Preliminaries

3.1 Problem Definition and Notations

In this paper, we assume that there are U users and V items. Following the convention of latent factor models, we represent each entity as a low-dimensional latent vector and the latent dimension is denoted as K . For a user u , we represent the latent assignment vector as ϕ_u , which indicates how the user will assign the budget to different item attributes. And we denote the latent rating vector as θ_u , which indicates how the user will rate each item attribute. So the orange bars in Figure 1 are illustrations of ϕ_u and θ_u . Meanwhile, the latent attribute vector of item v is denoted as β_v . The monetary allocation of the user u to item v is p_{uv} and the rating that the user u give to item v is y_{uv} .

We denote the observed feedback (the set of products that the user bought) of user u as O_u . For each unseen item v (the product that the user did not buy) of the user u , the task of a recommendation model m is to construct a ranking score

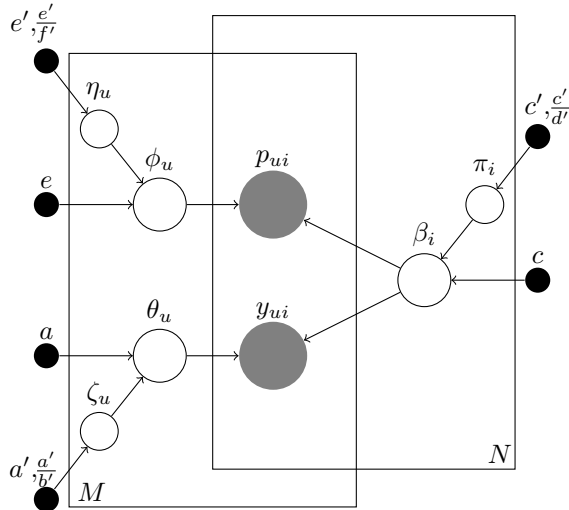


Figure 2: The graphical model of CBPF.

$f_u(v)$ based on O_u and then recommend the user with items with high ranking scores.

3.2 Hierarchical Poisson Factorization

Hierarchical Poisson factorization [Gopalan *et al.*, 2013] is a probabilistic collaborative filtering recommendation model for users' ratings. In hierarchical Poisson factorization, users and items are represented as low-dimensional and non-negative sparse vectors. The latent user vectors indicate user preferences and the latent item vectors indicate item attributes. The ratings that users give to items are assumed to be drawn from Poisson distributions. The rate of the Poisson distribution is set to be the inner product of the latent user vector and latent item vector. Meanwhile, each component of the latent user vector and the latent item vector is drawn from a Gamma distribution whose rate is assumed to be drawn from another Gamma distribution with fixed parameters.

4 Proposed Model

In CBPF, we try to uncover how users will assign their budgets to each item attribute and how users will rate the attribute accordingly. Unlike HPF [Gopalan *et al.*, 2013] which only considers users' ratings, we model users' ratings and budgets in a cohesive manner which captures the intuition that for intelligent recommendation it is important to involve users' budgets. The generative process of HPF is similar to CBPF and we omit it due to space limitation. Based on above analysis, the generative process of CBPF is defined as follows:

1. For each user u :

- (a) Sample $\zeta_u \sim \text{Gamma}(a', a'/b')$.
- (b) For each component k , sample preference

$$\theta_{uk} \sim \text{Gamma}(a, \zeta_u). \quad (1)$$

- (c) Sample $\pi_u \sim \text{Gamma}(e', e'/f')$.
- (d) For each component k , sample assignment

$$\phi_{uk} \sim \text{Gamma}(e, \pi_u). \quad (2)$$

2. For each item v :

- (a) Sample $\eta_u \sim \text{Gamma}(c', c'/d')$.
- (b) For each component k , sample attribute

$$\beta_{vk} \sim \text{Gamma}(c, \eta_u). \quad (3)$$

3. For each user u and item v ,

- (a) sample rating

$$y_{uv} \sim \text{Poisson}(\theta_u^T \beta_v). \quad (4)$$

- (b) sample allocation

$$p_{uv} \sim \text{Poisson}(\phi_u^T \beta_v). \quad (5)$$

In this paper, we use budget to refer to the monetary limit of a user, which equals to the sum of all allocations. The monetary allocation of the user u to item v is just the price of item v . And we use assignment to refer to the monetary allocation to a particular item attribute. So the latent assignment vector ϕ_u indicates how the user u will allocate her budget to different item attributes. The allocation p_{uv} of user u for item v is assumed to be drawn from a Poisson distribution whose rate is the inner product of the latent assignment vector and the latent attribute vector.

Users' ratings and allocations are highly correlated and both stem from the interaction of user preferences and product attributes. In CBPF, they are coupled in a natural and cohesive manner. HPF is a special case of the proposed model which only considers users' ratings. The latent variables of CBPF are $\theta_{1:U}, \phi_{1:U}, \beta_{1:V}, \zeta_{1:U}, \pi_{1:V}$ and $\eta_{1:U}$. The observed variables are users' ratings \mathbf{y} and users' allocations \mathbf{p} .

Interpretation: Due to the relation between Poisson distribution and multinomial distribution, we can re-write the generative process of users' allocations as a two-stage procedure similar to [Gopalan *et al.*, 2013]. The user u first decides an amount of budget b_u and then chooses how to allocate the budget across different products:

$$b_u \sim \text{Poisson}(\phi_u^T \sum_v \beta_v), \quad (6)$$

$$[p_{u1}, \dots, p_{uV}] \sim \text{Mult}(b_u, \frac{\phi_u^T \beta_v}{\phi_u^T \sum \beta_v}). \quad (7)$$

This interpretation makes it intuitive to model users' budgets with Poisson distribution. Moreover, the non-negative constraint of the latent assignment vectors makes it practically reasonable to model users' assignments of budgets.

Recommending new products: To recommend new products for each user, we fit the posterior distribution $p(\zeta_{1:U}, \theta_{1:U}, \pi_{1:V}, \beta_{1:V}, \eta_{1:U}, \phi_{1:U} | \mathbf{y}, \mathbf{p})$ using a variational inference algorithm [Jordan *et al.*, 1999; Blei *et al.*, 2016]. Once we fit the posterior, we can form recommendations for the user u by using the posterior expected Poisson parameters to compute the ranking score of each unseen item v of the user u ,

$$f_u(v) = \text{E}[(\theta_u^T + \phi_u^T) \beta_v | \mathbf{y}, \mathbf{p}] \quad (8)$$

The intuition of above recommendation score is that users' preferences consist of two parts: (1) how much budget the

user will allocate to the item, and (2) how the user will rate the item. We encode both the factors into above equation. The items with high ranking scores will be recommended to the user.

5 Approximate Posterior Inference

Given users’ purchasing history, our goal is to infer the latent rating vectors $\theta_{1:U}$, latent assignment vectors $\phi_{1:U}$, latent attribute vectors $\beta_{1:V}$ and the corresponding rate parameters. However, to compute the exact posterior distribution $p(\zeta_{1:U}, \theta_{1:U}, \pi_{1:V}, \beta_{1:V}, \eta_{1:U}, \phi_{1:U} | \mathbf{y}, \mathbf{p})$ is very difficult, so we adopt an approximate inference algorithm based on mean-field variational inference [Jordan *et al.*, 1999; Hoffman *et al.*, 2013; Blei *et al.*, 2016].

Variational inference framework first specifies a family of distributions Ω over the hidden variables and then approximates the exact posterior distribution by finding the best candidate from Ω which is closest to the exact posterior. In mean-field variational inference, each hidden variable is assumed to be independent from each other and is governed by its own variational parameters. This assumption simplifies the optimization procedure. We define the mean-field variational family as:

$$q(\zeta, \theta, \pi, \beta, \eta, \phi) = \prod_u q(\zeta_u | \tilde{\zeta}_u) \prod_{u,k} q(\theta_{uk} | \tilde{\theta}_{uk}) \prod_u q(\pi_u | \tilde{\pi}_u) \prod_{u,k} q(\beta_{uk} | \tilde{\beta}_{uk}) \prod_u q(\eta_u | \tilde{\eta}_u) \prod_{u,k} q(\phi_{uk} | \tilde{\phi}_{uk}). \tag{9}$$

This is a very flexible parameterization which can capture each user’s and each item’s unique characteristics. To measure the *closeness* of the variational distribution to the exact posterior, we adopt the Kullback-Leibler divergence. So the optimization task is as follows,

$$q^*(\zeta, \theta, \pi, \beta, \eta, \phi) = \underset{q \in \Omega}{\operatorname{argmin}} KL(q^*(\zeta, \theta, \pi, \beta, \eta, \phi) || p(\zeta, \theta, \pi, \beta, \eta, \phi | \mathbf{y}, \mathbf{p})). \tag{10}$$

The proposed model is one member of the important family of models named conditionally conjugate models [Blei *et al.*, 2016]. With the conditional conjugate property, we can easily obtain a closed-form coordinate ascent algorithm.

Conditionally conjugate models. Conditionally conjugate models are a family of models that the complete conditional of each hidden variable is in the exponential family and is in the same family as its prior. And the complete conditional is the conditional distribution of a hidden variable given the observations and other hidden variables in the model. The updates are obtained by setting each variational parameter equals to the expected natural parameter under the variational distribution q of the complete conditional.

Auxiliary variables. To obtain a fully conditionally conjugate model, we need to construct two auxiliary latent variable z^a and z^b . Following [Gopalan *et al.*, 2013], we define K latent variables $z_{uvk}^a \sim \text{Poisson}(\theta_{uk}\beta_{vk})$ such that $y_{uv} =$

$\sum_k z_{uvk}^a$ and K latent variables $z_{uvk}^b \sim \text{Poisson}(\phi_{uk}\beta_{vk})$ such that $p_{uv} = \sum_k z_{uvk}^b$. Recall that the sum of independent Poisson variables is a new Poisson variable. When marginalizing out these auxiliary variables, we preserve the marginal distributions of the observations.

Coordinate ascent algorithm. With the conditionals, we can easily develop a coordinate ascent variational algorithm. The update of each variational parameter is obtained by computing the expected natural parameter of the conditional under the variational distribution. We take the updates of parameters $\theta_{1:K}$ and z_{uv}^a as examples. The updates of other variational parameters follow similar patterns. The algorithm is shown in Algorithm 1.

The updates of the variational shape and rate parameter of the user preference vectors θ_{uk} are as follows,

$$\tilde{\theta}_{uk} = \langle a + \sum_v y_{uv} \tilde{z}_{uvk}^a, \frac{\zeta_u^{shp}}{\zeta_u^{rte}} + \sum_v \frac{\tilde{\beta}_{uk}^{shp}}{\tilde{\beta}_{uk}^{rte}} \rangle. \tag{11}$$

This equation is derived from the expected natural parameter of the complete conditional of θ_{uk} under the variational distribution q . We use $E_q[z_{uvk}^a] = y_{uv} \tilde{z}_{uvk}^a$ to update the shape parameter. And we use the expectation of a Gamma variable is the shape divided by the rate to update the rate parameter.

The update for the multinomial z_{uv}^a is

$$\tilde{z}_{uv}^a \propto \exp\{\Psi(\tilde{\theta}_{uk}^{shp}) - \log(\tilde{\theta}_{uk}^{rte}) + \Psi(\tilde{\beta}_{uk}^{shp}) - \log(\tilde{\beta}_{uk}^{rte})\}. \tag{12}$$

Where $\Psi()$ is the digamma function (the first derivative of the $\log \Gamma$ function). We use the fact that $E_q[\log \theta_{uk}] = \Psi(\tilde{\theta}_{uk}^{shp}) - \log \tilde{\theta}_{uk}^{rte}$. The hidden variables, complete conditionals and relevant expectations are shown in Table 1.

Algorithm 1 The variational inference algorithm of collaborative budget-aware Poisson factorization.

For all users and items, initialize the user parameters $\tilde{\theta}_u, \tilde{\phi}_u, \tilde{\zeta}_u^{rte}, \tilde{\pi}_u^{rte}$ and item parameters $\tilde{\beta}_v, \tilde{\eta}_v^{rte}$ randomly.

Set $\tilde{\zeta}_u^{shp}, \tilde{\pi}_u^{shp}$ and $\tilde{\eta}_v^{shp}$ as follows:

$$\tilde{\zeta}_u^{shp} = a' + Ka. \tilde{\pi}_u^{shp} = c' + Kc. \tilde{\eta}_v^{shp} = e' + Ke. \tag{13}$$

Repeat until convergence:

1. For each rating $r_{uv} > 0$, set \tilde{z}_{uv}^a to the expected conditional parameter of z_{uv}^a as shown in Table 1.
2. For each budget $p_{uv} > 0$, set \tilde{z}_{uv}^b to the expected conditional parameter of z_{uv}^b as shown in Table 1.
3. For each user u , update $\tilde{\theta}_u, \tilde{\zeta}_u^{rte}, \tilde{\phi}_u$ and $\tilde{\pi}_u^{rte}$ to their expected conditional parameters as shown in Table 1.
4. For each item v , update $\tilde{\beta}_v, \tilde{\pi}_v^{rte}$ to their expected conditional parameters as shown in Table 1.

6 Empirical Studies

6.1 Data Preparation

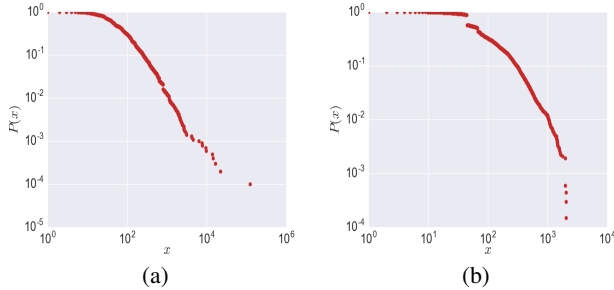
We test our model on several large datasets from Amazon.com [McAuley and Leskovec, 2013]. The datasets

Table 1: Hidden variables, complete conditionals, variational parameters and relevant expectations in CBPF.

Variable	Type	Complement Conditional	Parameters	Relevant Expectations
θ_{uk}	Gamma	$a + \sum_v z_{vik}^b, \pi_u + \sum_v \beta_{vk}$	$\tilde{\theta}_{uk}^{shp}, \tilde{\theta}_{uk}^{rte}$	$a + \sum_v y_{uv} z_{uvk}^a, \frac{\zeta_u^{shp}}{\zeta_u^{rte}} + \sum_v \frac{\tilde{\beta}_{uk}^{shp}}{\tilde{\beta}_{uk}^{rte}}$
ϕ_{uk}	Gamma	$e + \sum_v z_{vik}^a, \zeta_u + \sum_v \beta_{vk}$	$\tilde{\phi}_{uk}^{shp}, \tilde{\phi}_{uk}^{rte}$	$e + \sum_v p_{uv} z_{uvk}^b, \frac{\pi_u^{shp}}{\pi_u^{rte}} + \sum_v \frac{\tilde{\beta}_{uk}^{shp}}{\tilde{\beta}_{uk}^{rte}}$
β_{uk}	Gamma	$c + \sum_u z_{uvk}^a + \sum_u z_{uvk}^b, \eta_u + \sum_u \theta_{uk} + \sum_u \phi_{uk}$	$\tilde{\beta}_{uk}^{shp}, \tilde{\beta}_{uk}^{rte}$	$c + \sum_u y_{uv} z_{uvk}^a + \sum_u p_{uv} z_{uvk}^b, \frac{\eta_u^{shp}}{\eta_u^{rte}} + \sum_u \frac{\tilde{\theta}_{uk}^{shp}}{\tilde{\theta}_{uk}^{rte}} + \sum_u \frac{\tilde{\phi}_{uk}^{shp}}{\tilde{\phi}_{uk}^{rte}}$
ζ_u	Gamma	$a' + Ka, b' + \sum_k \theta_{uk}$	$\tilde{\zeta}_{uk}^{shp}, \tilde{\zeta}_{uk}^{rte}$	$a + \sum_v y_{uv} z_{uvk}^a, \frac{\zeta_u^{shp}}{\zeta_u^{rte}} + \sum_v \frac{\tilde{\beta}_{uk}^{shp}}{\tilde{\beta}_{uk}^{rte}}$
η_v	Gamma	$c' + Kc, d' + \sum_k \beta_{uk}$	$\tilde{\eta}_{uk}^{shp}, \tilde{\eta}_{uk}^{rte}$	$a + \sum_v y_{uv} z_{uvk}^a, \frac{\zeta_u^{shp}}{\zeta_u^{rte}} + \sum_v \frac{\tilde{\beta}_{uk}^{shp}}{\tilde{\beta}_{uk}^{rte}}$
π_u	Gamma	$e' + Ke, f' + \sum_k \phi_{uk}$	$\tilde{\pi}_{uk}^{shp}, \tilde{\pi}_{uk}^{rte}$	$a + \sum_v y_{uv} z_{uvk}^a, \frac{\zeta_u^{shp}}{\zeta_u^{rte}} + \sum_v \frac{\tilde{\beta}_{uk}^{shp}}{\tilde{\beta}_{uk}^{rte}}$
z_{uv}^a	Mult	$\log \theta_{uk} + \log \beta_{vk}$	z_{uv}^a	$\exp\{\Psi(\tilde{\theta}_{uk}^{shp}) - \log(\tilde{\theta}_{uk}^{rte}) + \Psi(\tilde{\beta}_{uk}^{shp}) - \log(\tilde{\beta}_{uk}^{rte})\}$
z_{uv}^b	Mult	$\log \phi_{uk} + \log \beta_{vk}$	z_{uv}^b	$\exp\{\Psi(\tilde{\phi}_{uk}^{shp}) - \log(\tilde{\phi}_{uk}^{rte}) + \Psi(\tilde{\beta}_{uk}^{shp}) - \log(\tilde{\beta}_{uk}^{rte})\}$

Table 2: Statistics of the Datasets.

	Shoes	Tools	Toys	Baby	Sports	Arts	Pet	Clothing	Cellphone	Patio
Users	12,421	4,566	8,869	7,057	14,124	1,815	2,035	14,560	1,485	2,700
Items	3,278	2,668	12,512	1,320	5,658	1,057	1,486	3,139	1,120	1,484
Observations	60,974	15,995	18,607	18,607	25,705	4,073	4,941	49,890	4,530	4,607
Average spending (\$)	416.30	11,985.25	243.94	129.69	122.83	71.60	61.26	68.72	170.26	183.83
Average rating	4.36	4.21	4.27	4.00	4.29	4.28	4.12	4.23	3.74	4.06


 Figure 3: Distributions of users' expenditures on products in the categories of *Sports* and *Baby*.

contain a variety of categories ranging from *Shoes* to *Cellphones*. Moreover, we crawled the price information of the products from a Amazon price tracking website ². The detailed statistics of the datasets are listed in Table 2. We are also interested in the diversities of users' total budgets. We plot the distributions of users' expenditures on products in the categories of *Sports* and *Baby* in Figure 2. The results show that few users spent much in online shopping while most of the users spent little. This phenomenon is in line with the user activity levels [Gopalan *et al.*, 2013].

We further normalize the prices to be compatible with the range of the ratings. We denote the maximum and minimum price of a specific category as p_{max} and p_{min} separately. And

²www.camelcamelcamel.com

the maximum rating is denoted as L , which is usually 5. We then translate a specific price p using $L \frac{p - p_{min}}{p_{max} - p_{min}}$ and round the result to the nearest integer.

6.2 Baselines

We compare our model with several state-of-the-art budget-unaware recommendation models.

HPF: Hierarchical Poisson factorization (HPF) [Gopalan *et al.*, 2013] is the state-of-the-art collaborative filtering recommendation model. It only considers users' ratings. We adopt the implementation of HPF by the original authors.

BPR: Bayesian personalized ranking (BPR) [Rendle *et al.*, 2009] is the state-of-the-art pairwise recommendation model for implicit feedback. We view ratings below 3 as positive feedback and implement BPR as described in [Rendle *et al.*, 2009].

CliMF: CliMF [Shi *et al.*, 2012] is the state-of-the-art ranking model for optimizing mean reciprocal rank. We use the implementation of CliMF of Graphchi [Kyrola *et al.*, 2012].

PopRank: This is a naive baseline which considers recommending the users with the most popular items in the training set.

6.3 Parameter Settings and Evaluations

Similar to the settings of [Gopalan *et al.*, 2013], we split each dataset into three parts: 70% of the dataset is served as training set, 20% of the dataset is served as test set and the remaining 10% is served as validation set. All the hyperparameters are tuned to work the best on the validation set. For CBPF and

Table 3: Results of precision-at-10 of the compared models.

	Shoes	Tools	Toys	Baby	Sports	Arts	Pet	Clothing	Cellphones	Patio
CBPF	0.12203	0.01614	0.05612	0.02930	0.03476	0.04653	0.01952	0.09549	0.02335	0.03364
HPF	0.11750	0.01508	0.04782	0.02457	0.03416	0.04365	0.01727	0.08509	0.02081	0.03225
BPR	0.10968	0.00801	0.01629	0.02229	0.01075	0.04790	0.01556	0.01535	0.02284	0.02469
ClIMF	0.02515	0.00396	0.00461	0.02707	0.01197	0.01942	0.01209	0.01124	0.01838	0.01926
PopRank	0.00150	0.00362	0.00475	0.02733	0.00886	0.01813	0.01266	0.01189	0.02017	0.01802

Table 4: Results of NCCR of the compared models.

	Shoes	Tools	Toys	Baby	Sports	Arts	Pet	Clothing	Cellphones	Patio
CBPF	0.32247	0.11307	0.09571	0.16953	0.08610	0.13842	0.04145	0.21608	0.15588	0.16473
HPF	0.30376	0.10845	0.08603	0.15072	0.07968	0.11528	0.03265	0.20645	0.15271	0.15964
BPR	0.02342	0.01345	0.02105	0.08304	0.01980	0.06993	0.03113	0.05042	0.03766	0.03852
ClIMF	0.03102	0.03526	0.00729	0.04257	0.02823	0.07507	0.03061	0.04464	0.14778	0.14759
PopRank	0.00269	0.01828	0.00573	0.04919	0.00984	0.03470	0.02741	0.02118	0.04348	0.03081

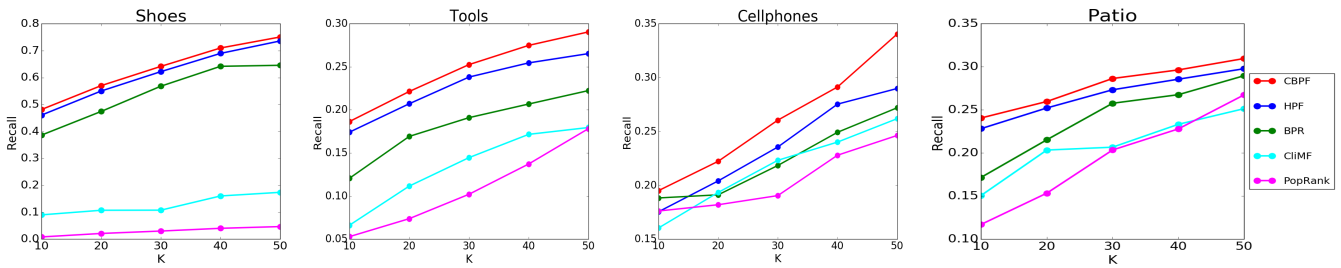


Figure 4: Results of recall-at- K of the compared models.

HPF, we set each Gamma shape and rate hyperparameter to 0.3. For BPR, the learning rate is 0.001 and the regularization parameter is 0.5. For ClIMF, the learning rate is 0.005 and the regularization parameter is 0.01. And we fix the dimension of the latent vectors of all models to 10 for fair comparison.

In each round, we compute the average predictive log-likelihood of the validation set, the convergence of CBPF is guaranteed when the change of the log-likelihood is less than 0.001%. After fitting the model in the training set, for a specific user u , we calculate the ranking score $f_u(v)$ for each item v in the test set. The items with higher ranking scores are ranked higher in the recommended list. We adopt precision-at- K , recall-at- K and normalized cumulative reciprocal rank (NCCR) [Chaney *et al.*, 2015] to measure the ranking performance of the models.

6.4 Results Analysis

We show the results of precision-at-10 and NCCR of all models in Table 3 and Table 4. Best performance is in bold font. And the results of recall-at- K ($K \in \{10,20,30,40,50\}$) are shown in Figure 4. As we can see, CBPF nearly beats all the baselines across all the categories. In most of the categories the improvements are significant. Moreover, our model exhibits particularly obvious improvements over some *cold start* categories where users have few ratings. The average improvement of precision-at-10 is 9.72%. The results empirically prove the importance of modeling users' budgets

for recommendation. By finding what attributes the users will assign more budgets and avoiding recommending the users with inappropriately expensive or cheap products, we can form more reasonable and more accurate recommendation results. The performance of BPR and ClIMF is consistently lower than CBPF and HPF. One plausible reason is that their optimization goal is not directly related to the ranking metrics we compare here. We also examine the parameter sensitivity of CBPF. The results show that our model is rather stable with respect to the change of hyperparameters.

7 Conclusion

In this paper, we propose an original generative model named collaborative budget-aware Poisson factorization (CBPF) to connect users' ratings and budgets in a natural and cohesive manner. The model enjoys high interpretability and emphasizes the significance of modeling users' budgets. Extensive experimental results empirically prove the rationality and importance of the proposed model. We hope that the proposed model will prove useful for online shopping websites. Future works include considering modeling users' ratings and budgets in a dynamic and sequential manner.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (NSFC) No.61672449 and No.61472347.

References

- [Blei *et al.*, 2016] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *arXiv preprint arXiv:1601.00670*, 2016.
- [Chaney *et al.*, 2015] Allison JB Chaney, David M Blei, and Tina Eliassi-Rad. A probabilistic model for using social networks in personalized item recommendation. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 43–50. ACM, 2015.
- [Charlin *et al.*, 2015] Laurent Charlin, Rajesh Ranganath, James McInerney, and David M Blei. Dynamic poisson factorization. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 155–162. ACM, 2015.
- [Du and Kamakura, 2008] Rex Y Du and Wagner A Kamakura. Where did all that money go? understanding how consumers allocate their consumption budget. *Journal of Marketing*, 72(6):109–131, 2008.
- [Goldberg *et al.*, 1992] David Goldberg, David Nichols, Brian M Oki, and Douglas Terry. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12):61–70, 1992.
- [Gopalan *et al.*, 2013] Prem Gopalan, Jake M Hofman, and David M Blei. Scalable recommendation with poisson factorization. *arXiv preprint arXiv:1311.1704*, 2013.
- [Gopalan *et al.*, 2014] Prem K Gopalan, Laurent Charlin, and David Blei. Content-based recommendations with poisson factorization. In *Advances in Neural Information Processing Systems*, pages 3176–3184, 2014.
- [Gourinchas and Parker, 2002] Pierre-Olivier Gourinchas and Jonathan A Parker. Consumption over the life cycle. *Econometrica*, 70(1):47–89, 2002.
- [Hoffman *et al.*, 2013] Matthew D Hoffman, David M Blei, Chong Wang, and John William Paisley. Stochastic variational inference. *Journal of Machine Learning Research*, 14(1):1303–1347, 2013.
- [Jordan *et al.*, 1999] Michael I Jordan, Zoubin Ghahramani, Tommi S Jaakkola, and Lawrence K Saul. An introduction to variational methods for graphical models. *Machine learning*, 37(2):183–233, 1999.
- [Kao *et al.*, 2001] Chihwa Kao, Lung-fei Lee, Mark M Pitt, et al. Simulated maximum likelihood estimation of the linear expenditure system with binding non-negativity constraints. *Annals of Economics and Finance*, 2(1):203–223, 2001.
- [Kooti *et al.*, 2016] Farshad Kooti, Kristina Lerman, Luca Maria Aiello, Mihajlo Grbovic, Nemanja Djuric, and Vladan Radosavljevic. Portrait of an online shopper: Understanding and predicting consumer behavior. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 205–214. ACM, 2016.
- [Koren and Bell, 2011] Yehuda Koren and Robert Bell. Advances in collaborative filtering. In *Recommender systems handbook*, pages 145–186. Springer, 2011.
- [Koren *et al.*, 2009] Yehuda Koren, Robert Bell, Chris Volinsky, et al. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.
- [Koren, 2008] Yehuda Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 426–434. ACM, 2008.
- [Kyrola *et al.*, 2012] Aapo Kyrola, Guy Blelloch, and Carlos Guestrin. Graphchi: large-scale graph computation on just a pc. In *Presented as part of the 10th USENIX Symposium on Operating Systems Design and Implementation (OSDI 12)*, pages 31–46, 2012.
- [McAuley and Leskovec, 2013] Julian McAuley and Jure Leskovec. Hidden factors and hidden topics: understanding rating dimensions with review text. In *Proceedings of the 7th ACM Conference on Recommender Systems*, pages 165–172. ACM, 2013.
- [Rendle *et al.*, 2009] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*, pages 452–461. AUAI Press, 2009.
- [Resnick and Varian, 1997] Paul Resnick and Hal R Varian. Recommender systems. *Communications of the ACM*, 40(3):56–58, 1997.
- [Salakhutdinov and Mnih, 2008a] Ruslan Salakhutdinov and Andriy Mnih. Bayesian probabilistic matrix factorization using markov chain monte carlo. In *Proceedings of the 25th International Conference on Machine Learning*, pages 880–887. ACM, 2008.
- [Salakhutdinov and Mnih, 2008b] Ruslan Salakhutdinov and Andriy Mnih. Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems*, volume 20, 2008.
- [Shi *et al.*, 2012] Yue Shi, Alexandros Karatzoglou, Linas Baltrunas, Martha Larson, Nuria Oliver, and Alan Hanjalic. Clmf: learning to maximize reciprocal rank with collaborative less-is-more filtering. In *Proceedings of the sixth ACM conference on Recommender systems*, pages 139–146. ACM, 2012.
- [Xie *et al.*, 2010] Min Xie, Laks VS Lakshmanan, and Peter T Wood. Breaking out of the box of recommendations: from items to packages. In *Proceedings of the fourth ACM conference on Recommender systems*, pages 151–158. ACM, 2010.