A Joint Model for Who-to-Follow and What-to-View Recommendations on Behance

Maja R. Rudolph
Department of Computer Science, Columbia University
maja@cs.columbia.edu

Matthew Hoffman
Creative Technologies Lab
Adobe Systems
mathoffm@adobe.com

Aaron Hertzmann
Creative Technologies Lab
Adobe Systems
hertzman@adobe.com

ABSTRACT

Good recommendations are a key tool to increase user engagement and user satisfaction on many social networks. Here we focus on Behance, a social network for artist from various fields such as typography, street art, industrial design, and fashion. On Behance, the artists can connect by following each other, display their work in online portfolios, and brows each other's work. Each user has a personalized dashboard which is an integral part of the Behance experience.

In this work we create a joint behavior model which jointly models the users' viewing behavior and the social network. The joint model which we fit with variational inference is capable of producing both who-to-follow and what-to-view recommendations. We show on real data from Behance that the joint behavior model outperforms a Poisson factorization approach which treats both data sources separately.

1. INTRODUCTION

Many web applications thrive on an engaged and happy userbase. Often the key to keeping users happy and engaged is to give them good personalized recommendations as to which content to look at, which news articles to read, what music they might want to listen to, what product to buy, what movie to watch, which old friend to connect with.

Many recommendation systems use past user behavior data such as what a user has previously looked at to find new items a user might be interested in. Collaborative filtering approaches such as matrix factorization [11, 12] and Poisson factorization for count data [2] embed both the users and the items in a latent abstract space which captures similarities between users and items based on the past behavior data. Users with similar past behavior will be embedded closely to each other and their similarities are used to extrapolate recommendations.

In this work we propose a probabilistic model for modeling two types of user behavior data together. We show that by combining viewing behavior data with following behavior data we can improve recommendations on Behance.

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
http://dx.doi.org/10.1145/2872518.2890083.

The Behance social network. Behance is a social network for artists, graphic designers, and creatives from various fields such as typography, street art, industrial design and fashion. On Behance, the artists can share their projects in digital portfolios and browse each other's work. Many users choose to subscribe to artists whose work they are interested in by following them. Each user has a personalized dashboard with suggested portfolios to view. The dashboard displays the portfolios of recent projects of followed artists.

Hence, users who follow many artists whose work they are very interested in will find the dashboard more useful for browsing artwork than users who follow few or irrelevant artists. Hence, good who-to-follow recommendations are important to improving a users' dashboard. In addition, the dashboard can be improved by reranking the portfolios on the dashboard according to what-to-view recommendations of the system.

What-to-view and who-to-follow recommendations. A typical collaborative filtering approach addresses the what-to-view and the who-to-follow recommendation task separately from each other. For what-to-view recommendations we can learn the users' viewing preferences from the data of which portfolios the users have previously looked at, i.e. the matrix of user-portfolio viewing behavior. The who-to-follow recommendations would be handled by a separate model which extracts the user' following preferences from a user-user following matrix. We present a Poisson factorization approach for both recommendation tasks in Sec. 2.1.

A joint behavior model. In this work we show that by using a unified model for both data sources we can improve both the who-to-follow and the what-to-view recommendations. The following reasons support using a joint model:

1. **More data to learn the users' preferences:** Both information on what portfolios a user has viewed and who the user has followed sheds light on what kind of artwork/artists the user likes. By combining both data sources we have more data to learn the users' preferences.

2. **Confounding by social structure:** The social network (i.e. the following behavior of the users) directly affects the users' browsing behavior on Behance. Many users will look at the projects posted by a close friend (or their boss) simply because they know the person and not necessarily because the work exactly matches their artistic interests and preferences. By explicitly modeling the effect of the social network on the viewing behavior we can gain a more granular understanding of the preferences of the users and avoid misinterpreting interest for social reasons as interest because of viewing preference.

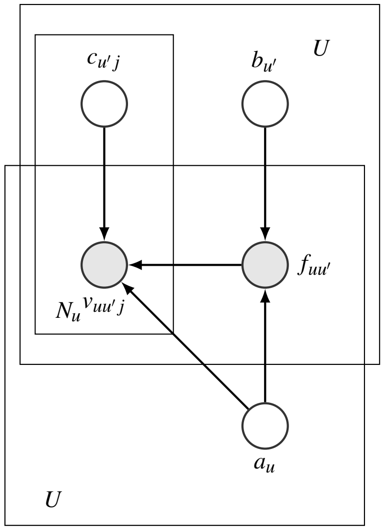

Figure 1: Joint model for viewing and following behavior of Behance users.

3. **Improved recommendation performance:** As we show in Sec. 3 on real data from Behance the joint model outperforms the models for the individual data sources both in terms of who-to-follow and what-to-view recommendations.

2. METHOD

Consider N Behance users, each of which has N_u projects publicly posted on Behance. Next assume we observed a list $\mathscr{F}$ of user user pairs (u, u') denoting that user u is following user u'. For notation we introduce binary variables $f_{uu'}$ which indicate whether user u follows user u' or not. In addition, we observe a list $\mathscr{V}$ of user user item triplets (u, u', j) denoting that user u has viewed project j of user u'. Accordingly, the binary variable $v_{uu'j}$ is an indicator whether user u has viewed project j of user u' or not.

We first present Poisson factorization for separately modeling $\mathscr{F}$ and $\mathscr{V}$. The joint behavior model is then presented in Sec. 2.2.

2.1 Poisson Factorization

Poisson factorization is a collaborative filtering model for count data. Here we review the model that uses the viewing data $\mathscr{V}$ to embed each user u and each project $u'j$[1] into a latent space which captures the similarity between users and their affinity to like other projects they have not viewed before. In the latent space, two users' locations will be close to each other, when they have had similar past viewing behavior. The projects are also embedded in the same space and their latent locations will be close to each other when they have been viewed by many users with similar preferences.

The dimensionality of the embedding K, has to be fixed in advance. Each user u is associated with a latent preference vector $a_u \in \mathbb{R}^K$ and each project $u'j$ is associated with a latent style attribute vector $c_{u'j} \in \mathbb{R}^K$.

The generative process is the following:

1. For each user $u \in \{1, \cdots, U\}$ draw each component $k \in \{1, \cdots, K\}$ from a gamma distribution:

$$a_{u,k} \sim \text{Gamma}(\alpha_{\text{shp}}, \alpha_{\text{rte}}) \tag{1}$$

[1] we use and indexing strategy for the projects which also tracks the user u' which owns project j. This convention is useful for the joint model presented in the next section.

2. For each user $u' \in \{1, \cdots, U\}$, for each project $j \in \{1, \cdots, N_u\}$, draw each component $k \in \{1, \cdots, K\}$ from a gamma distribution:

$$c_{u'j,k} \sim \text{Gamma}(\gamma_{\text{shp}}, \gamma_{\text{rte}}) \tag{2}$$

3. For each triplet (u, u', j), draw the observation from a Poisson likelihood:

$$v_{uu'j} \sim \text{Poisson}(a_u^T c_{u'j}) \tag{3}$$

A similar Poisson factorization model can be used to model the following behavior. There each user u is still associated with a latent preference vector a_u but also with a latent creative skill vector $b_{u'}$. The mean of the (Poisson) likelihood of user u following user u' is now $a_u^T b_{u'}$.

In this paper we propose a model which builds on Poisson factorization but jointly models the viewing and following behavior. The joint behavior model is presented in the next section.

2.2 Joint Behavior Model

Like Poisson factorization, the joint behavior model embeds the users and the items in a latent space which captures similarities between users in terms of their follow and viewing preferences and in terms of their creative skills and captures the similarities between projects in terms of their latent style attributes. Each user has two locations in the latent space. One for their artistic skills and one for their preferences. The preference vector of a user will affect both which artists the user is likely to follow and which projects the user is likely to view. This means that during inference for the joint behavior model both data sources can be used to learn a user's preferences. Having the additional data is useful in many ways but is particularly exemplified by trying to learn the preferences of a user who has browsed some projects but not yet followed anyone. The Poisson model from the previous section cannot induce any estimate of who a user might want to follow before this user has decided to follow at least one person. The joint model we propose here, will facilitate estimating a users preferences for following someone from the user's project viewing behavior.

The joint behavior model also addresses that users view projects for different reasons which do not necessarily reflect only their viewing preferences. In this work we explicitly model that some views happen because the user follows the owner of the project. The likelihood of user u viewing project $u'j$ owned by user u' is a weighted average of different explanations. Either user u does not follow user u' but genuinely is interested in user u''s work, or user u does follow user u'. In that case, user u is either genuinely interested in the project or user u is viewing it just because user u' is the owner.

Imagine for example that user u is a graphic designer. In the following three scenarios different "explanations" would dominate the likelihood term of whether user u viewed project $u'j$. When u' is a random artist user u does not know then user u's latent preferences together with the style attributes of project $u'j$ affect the view likelihood most. When user u' is someone user u follows on Behance the follow might affect whether the viewing likelihood in one of two ways. User u' could be a fantastic fellow graphic designer whose work user u genuinely likes to brows for inspiration. On the other hand, user u' could be a relative or friend of user u and so u is browsing their work not because user u genuinely likes it but for social reasons.

The generative process of the joint behavior model is as follows:

Table 1: Results on the who-to-follow and what-to-view recommendation task. The joint behavior model outperforms the separately trained Poisson factorization models in terms of precision and recall at 20 on both tasks.

		JBM (K=10)	PFF (K=10)	PFV (K=10)
who-to-follow	precision @20	**30.01 +- 0.80**	23.91 +- 0.69	-
recommendations	recall @20	**17.53 +- 0.42**	13.85+-0.35	-
what-to-view	precision @20	**17.58 +- 0.45**	-	12.42 +- 0.38
recommendations	recall @20	**2.98 +- 0.09**	-	1.77 +- 0.05

1. For each user $u \in \{1, \cdots, U\}$ draw each component $k \in \{1, \cdots, K\}$ of the user's preference vector from a gamma distribution:

$$a_{u,k} \sim \text{Gamma}(\alpha_{\text{shp}}, \alpha_{\text{rte}}) \tag{4}$$

2. For each user $u \in \{1, \cdots, U\}$ draw each component $k \in \{1, \cdots, K\}$ of the user's artistic skill vector from a gamma distribution:

$$b_{u,k} \sim \text{Gamma}(\beta_{\text{shp}}, \beta_{\text{rte}}) \tag{5}$$

3. For each user $u' \in \{1, \cdots, U\}$, for each project $j \in \{1, \cdots, N_u\}$, draw each component $k \in \{1, \cdots, K\}$ from a gamma distribution:

$$c_{u'j,k} \sim \text{Gamma}(\gamma_{\text{shp}}, \gamma_{\text{rte}}) \tag{6}$$

4. Draw the scaling parameters

$$f_0, v_0, v_1, v_2, v_3 \sim \text{Gamma}(\sigma_{\text{shp}}, \sigma_{\text{rte}}) \tag{7}$$

5. For each pair (u, u') with $u \neq u'$, draw the observation whether u follows u' from a Poisson likelihood:

$$f_{uu'} \sim \text{Poisson}(f_0 + a_u^T b_{u'}) \tag{8}$$

6. For each triplet (u, u', j), draw the observation from a Poisson likelihood:

$$v_{uu'j} \sim \text{Poisson}(v_0 + v_1 a_u^T c_{u'j} + v_2 f_{uu'} + v_3 f_{uu'} a_u^T c_{u'j}) \tag{9}$$

2.3 Variational Inference

The next goal is to use the observed following and view data to estimate the user preferences, their artistic skills and the project specific attributes that best explain the data under the joint behavior model.

To get these estimates we want to compute the posterior distribution $p(\theta|\mathscr{V}, \mathscr{F})$ over the model parameters $\theta = \{a_{u,k}, b_{u,k}, c_{u'j,k}, f_0, v_{0:3}\}$. Since this computation is intractable we use the mean-field assumption and approximate the posterior with a fully factored variational distribution [9, 14], which requires that we define the functional forms of the q distributions. The parameters of the q distributions are the variational parameters. We place a gamma distribution on all parameters in θ whose shape and rate parameter we need to optimize. In addition we introduce multinomial auxiliary variables as in [2].

The general variational objective we seek to maximize with respect to the variational parameters of a particular q is

$$\mathscr{L} = \mathbb{E}_q[\log p(\mathscr{V}, \mathscr{F}, \theta)] - \mathbb{E}_q[\log q(\theta)]. \tag{10}$$

which can be shown to minimize the Kullback-Leibler divergence between $q(\theta)$ and the true posterior $p(\theta|\mathscr{F}, \mathscr{V})$. Using the standard procedure, each $q(\theta_i)$ can be computed using the formula

$$q(\theta_i) \propto \exp\{\mathbb{E}_{q(\theta_{-i})}[\log p(\mathscr{F}, \mathscr{V}, \theta)]\}. \tag{11}$$

Variational inference turns the posterior inference task into an optimization task. In each iteration we fix all but one latent variable and optimize the variational parameters of it's q-distribution using Eqn. 11. We run this coordinate ascent algorithm until the variational objective (Eqn.10) stops increasing.

3. EMPIRICAL STUDY

We study the joint behavior model on a data set of real user interactions on Behance. The data set contains about $1M$ follows and $16M$ views of $100K$ unique users with a total of $400K$ portfolios. In around 20% of the views of this data set, the viewing user also follows the owner of the portfolio. The other 80% of the views are not associated with a follow.

We fit three models to the data:

- *Poisson factorization on the follows (PFF)*
- *Poisson factorization on the views (PFV)*
- *Joint behavior model on the follows and views (JBM)*

We compare JBM to Poisson factorization on two recommendation tasks: who-to-follow and what-to-view recommendations.

The data $\mathscr{F}, \mathscr{V}$ are randomly split into train and test data as follows: For each user u we randomly select 10% of the users user u follows and ignore the follow in the training data $\mathscr{F}_{\text{train}}$. Instead this will be part of the test data $\mathscr{F}_{\text{test}}$.

We also make all the views user u has made of projects the held-out users own part of the test data $\mathscr{V}_{\text{test}}$ and ignore these views in the training data $\mathscr{V}_{\text{train}}$.

Additionally, we hold out another 10% of the views of each user at random. These held-out views are part of $\mathscr{V}_{\text{test}}$ but not of $\mathscr{V}_{\text{train}}$.

Results are reported in terms of precision and recall at 20 in Table 1. The joint behavior model significantly outperforms Poisson factorization in both tasks.

4. DISCUSSION

We have presented a probabilistic model for both the viewing behavior and social network of users on Behance. In the joint model, the latent preferences of a user affect both what kind of artists the user might want to follow as well as the user's interest in the art projects on Behance. Because of this shared mechanism, both the following behavior data and the viewing behavior data sheds light on a users preferences.

This can help with the cold start problem in some cases: When a user has not followed any users yet the unimodal Poisson factorization model would have no data to provide who-to-follow recommendations even if this user has already browsed many projects. This is not an issue for the joint model which can already extract some information about the users' preferences for who they might want to follow from their viewing behavior.

The model can be fit efficiently using variational inference (Sec. 2.3) and in Sec. 3 we empirically show that the joint model improves who-to-follow and what-to-view recommendations which are important for increasing user engagement on Behance.

References

[1] K. Chen, T. Chen, G. Zheng, O. Jin, E. Yao, and Y. Yu. Collaborative personalized tweet recommendation. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 661–670. ACM, 2012.

[2] P. Gopalan, J. M. Hofman, and D. M. Blei. Scalable recommendation with poisson factorization. *arXiv preprint arXiv:1311.1704*, 2013.

[3] P. Gopalan, F. J. Ruiz, R. Ranganath, and D. M. Blei. Bayesian nonparametric poisson factorization for recommendation systems. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, pages 275–283, 2014.

[4] P. K. Gopalan, L. Charlin, and D. Blei. Content-based recommendations with poisson factorization. In *Advances in Neural Information Processing Systems*, pages 3176–3184, 2014.

[5] P. Gupta, A. Goel, J. Lin, A. Sharma, D. Wang, and R. Zadeh. Wtf: The who to follow service at twitter. In *Proceedings of the 22nd international conference on World Wide Web*, pages 505–514. International World Wide Web Conferences Steering Committee, 2013.

[6] M. D. Hoffman, D. M. Blei, C. Wang, and J. Paisley. Stochastic variational inference. *The Journal of Machine Learning Research*, 14(1):1303–1347, 2013.

[7] D. Hu and T. Schneiter. Targeted content for a real-time activity feed: For first time visitors to power users. In *Proceedings of the 24th International Conference on World Wide Web Companion*, pages 1269–1274. International World Wide Web Conferences Steering Committee, 2015.

[8] D. J. Hu, R. Hall, and J. Attenberg. Style in the long tail: Discovering unique interests with latent variable models in large scale social e-commerce. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1640–1649. ACM, 2014.

[9] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul. An introduction to variational methods for graphical models. *Machine learning*, 37(2):183–233, 1999.

[10] K. Y. Kamath, A.-M. Popescu, and J. Caverlee. Board recommendation in pinterest. In *UMAP Workshops*, 2013.

[11] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[12] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In *NIPS*, volume 1, pages 2–1, 2007.

[13] R. Salakhutdinov and A. 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.

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

[15] J. Zhang, B. Liu, J. Tang, T. Chen, and J. Li. Social influence locality for modeling retweeting behaviors. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, pages 2761–2767. AAAI Press, 2013.

[16] C. Zhong, D. Karamshuk, and N. Sastry. Predicting pinterest: Automating a distributed human computation. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1417–1426. International World Wide Web Conferences Steering Committee, 2015.