Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

Recommending Groups to Users Using User-Group Engagement and Time-Dependent Matrix Factorization

Xin Wang[†‡], Roger Donaldson[♮], Christopher Nell[§], Peter Gorniak[§], Martin Ester[‡] and Jiajun Bu[†]
[†]Zhejiang Provincial Key Laboratory of Service Robot, College of Computer Science, Zhejiang University, China
[‡]School of Computing Science, Simon Fraser University, Burnaby, B.C., Canada
[♮]Department of Mathematics, The University of British Columbia, Vancouver, B.C., Canada
[§]DeviantArt, Inc. Vancouver, B.C., Canada
[‡]{xwa49,ester}@cs.sfu.ca,[†]{xinwang,bjj}@zju.edu.cn,[♮]rdonald@math.ubc.ca,[§]{c.nell,pgorniak}@deviantart.com

Abstract

Social networks often provide group features to help users with similar interests associate and consume content together. Recommending groups to users poses challenges due to their complex relationship: user-group affinity is typically measured implicitly and varies with time; similarly, group characteristics change as users join and leave. To tackle these challenges, we adapt existing matrix factorization techniques to learn user-group affinity based on two different implicit engagement metrics: (i) which group-provided content users consume; and (ii) which content users provide to groups. To capture the temporally extended nature of group engagement we implement a time-varying factorization. We test the assertion that latent preferences for groups and users are sparse in investigating elastic-net regularization. Our experiments indicate that the time-varying implicit engagement-based model provides the best top-K group recommendations, illustrating the benefit of the added model complexity.

Introduction

Online web services recommend content items such as music, movies, or books etc. to users via algorithmic recommendations. Many of these recommenders are based on the principle of collaborative filtering, suggesting items that similar users have consumed. On the other hand, more and more social media and consumer websites are providing mechanisms by which users can self-organize into groups with other users having similar opinions or interests.

The problem of recommending groups to users has been investigated in the literature. In particular, methods for factorizing the user-group membership matrix have been proposed, using group features, user-item ratings and user-user networks to improve the performance. However, the existing methods fail to capture the dynamics of user group relationships: while the properties of content items typically do not change in the course of time, groups tend to evolve as users join or leave. Moreover, the preferences of users themselves tend to change over time. Thus, the existing methods are not completely adequate for recommending groups to users.

In this paper we focus on the problem of recommending groups to users using implicit measures of user-group affinity, and model the time-varying nature of such measures.

To quantify user-group affinity, we define a user-group engagement matrix that is constructed from more data than simple Boolean user-group membership information. Rather than relying on explicit surveys of user ratings of groups which are not always available, a user's affinity for a group is measured implicitly through observing how often and in what manner the user engages with that group. Following (Hu, Koren, and Volinsky 2008), our confidence in a user's affinity for a group increases with the number of interactions between the user and the group.

Unlike many types of content where user-item engagement occurs within a relatively brief consumption period, user-group interaction occurs over extended time scales. For example, a user may engage with a particular group 10 times in one month, 20 times the next, but only 2 times in the third month, and not at all in the fourth. This kind of extended interaction offers an opportunity to explicitly model changes in both user preferences and group dynamics when producing recommendations. Our intuition is that doing so will lead to improved recommendations.

We capture the time-varying nature of the group recommendation problem in two ways. First, we propose a time-varying matrix factorization in order to capture preference changes. Second, we introduce two time-varying user biases and one time-varying group bias in order to capture changes in activity levels. We incorporate time series analysis to model the temporal evolution of these factors and biases. Thus, our work makes three main contributions:

- When recommending groups to users, we consider not only group membership, but also different kinds of engagement between users and groups.
- We model evolution of the preferences and activity levels of both users and groups in order to better predict future preferences.
- We evaluate our methods using three real-world datasets from DeviantArt (DeviantArt), a large social network for artists and art enthusiasts. Our experiments show that using implicit engagement measures instead of Boolean membership improves recommendation performance. Taking the temporal nature of the engagement into account produces further improvements, and we also see a moderate improvement from the use of non-negative factorization with elastic-net regularization.

Related Work

Temporal recommendation There has been some work on temporal recommendation of items to users. Koren (Koren 2010) combines collaborative filtering and temporal dynamics together by proposing a model tracking the temporal evolution of user behaviour throughout the life span of the items. Other authors have examined time-dependent methods such as tensor factorization, session-based temporal graph model and dynamic matrix factorization etc. (Xiong et al. 2010; Xiang et al. 2010; Koenigstein, Dror, and Koren 2011; Lathia, Hailes, and Capra 2009; Chua, Oentaryo, and Lim 2013). This work captures past temporal patterns, but does not extrapolate future temporal dynamics to estimate future changes in users' preferences. Zhang et al. (Zhang et al. 2014) incorporate a transition matrix into conventional (Salakhutdinov and Mnih 2008b) and Bayesian (Salakhutdinov and Mnih 2008a) probabilistic matrix factorization methods, modelling the evolution of user preferences under the assumption that future preferences depend only on the immediately preceding state. Our work relaxes this assumption, instead assuming that both user and group characteristics change smoothly over time and predicting future preferences based on the complete interaction history. We note that conceptually Zhang's method is a special case of our method when we remove biases from our method, assume future user preferences depend only on one instead of p preceding states, and assume group characteristics remain unchanged in the course of time.

Group recommendation The problem of recommending groups to users has been studied elsewhere (Chen et al. 2009; Chen, Zhang, and Chang 2008; Vasuki et al. 2010; Wang et al. 2012; Zeng and Chen 2013) through exploring a variety of probabilistic and combinatorial recommendation methods applied to Boolean user-group membership matrices, with increasing success as more side information is incorporated into the model. Specifically, Chen (2008) proposes incorporating a probabilistic model which also considers the group-word matrix derived from the textual description of the group, and Vasuki (2010) takes an additional user-user linkage matrix derived from social network relationships into consideration. Zeng (2013) incorporates both user-item ratings and user-user social relationships (called "heterogeneous resources") into the user-group membership matrix. However, none of these authors model the temporal dynamics of user-group engagement. For completeness, in literature *group recommendation* also refers to the problem of recommending items to a group of users, which is not our focus in this paper and we refer readers to (Ali and Kim 2015; Baltrunas, Makcinskas, and Ricci 2010) for more detailed information if interested.

User-Group Engagement

In this section, we introduce the group features on our model data source, DeviantArt, as well as our proposed user-group engagement measures.

DeviantArt (DeviantArt) is the world's largest online arts community, with more than 60 million unique monthly visitors and 30 million registered users. Approximately 10,000 new users register daily. DeviantArt users submit over 150,000 new artworks every day, and the site has received over 300 million submissions in total. Art appreciators can engage with art by *favouriting* artwork (2 million events/day), or by *commenting* on artwork (also 2 million events/day). Groups on DeviantArt are self-organized associations of users who have the ability collectively curate art. This collective curation provides value to artists, who benefit from the endorsement provided by well-known groups, and to individual art collectors, who can use these curated collections to discover new art.

Measuring User-Group Engagement

We consider two variants of the user-group recommendation problem: recommending groups to artists, and recommending groups to collectors. Since DeviantArt users do not explicitly identify themselves as either artists or collectors, and since DeviantArt provides no mechanism for explicit rating of groups, we instead measure artist-group and collector-group affinity implicitly, via user-group engagement. On DeviantArt, artists provide new artwork to their groups, and collectors consume artwork provided by their groups. Therefore, we define two types of user-group engagement:

production engagement: the number of art submissions by user u that were accepted by group g

consumption engagement: the number of art submissions accepted by group g that user u subsequently favourited

We might use production engagement when recommending groups to artists, and consumption engagement when recommending groups to collectors.

Although the details of the definitions of production and consumption engagement above are specific to DeviantArt, we argue that the high level concepts of these definitions transfer to other scenarios of user groups. For example, Douban (Douban), a Chinese online social network, allows its users to create content related to films, books, music, and recent events and activities in their self-organized groups (called Douban Group). Users engage with their groups through creating new content such as articles or polls (production engagement) and through reading articles or participating in a poll (consumption engagement). Generally, the production engagement measures which content users provide to groups and the consumption engagement measures which group-provided content users consume.

Compared to Boolean membership or explicit item ratings data, the user-group engagement exhibits three interesting properties:

1. User-group engagement is nonnegative but otherwise unbounded; explicit ratings are usually restricted to a closed interval (e.g., integers from 1 to 5).
2. Users engage with groups gradually over extended time periods; user-item ratings are typically collected at a single time point.
3. Group characteristics change over time (for example, as users join and leave and activity levels increase or decrease); item characteristics typically do not.

The proposed recommendation method is designed to exploit each of these three properties.

Recommending Groups to Users

To the best of our knowledge, all the existing online social websites offering group features provide no formal mechanism for users to rate groups, which motivates our use of the matrix factorization based on an implicit feedback scheme as our starting point. Our implicit scheme incorporates the strength of user-group interactions, which we measure as production and consumption engagement. As group members' tastes change, user preferences and group properties tend to subsequently change over time. Hence, after introducing a static model, we propose a temporal model intended to capture this time-dependence.

Static Model

Hu et al. (Hu, Koren, and Volinsky 2008) predict users' preferences for TV programs through an implicit scoring model whose factors are computed by the matrix factorization

$$\begin{aligned} \underset{\mathbf{X}_u,\mathbf{Y}_i}{\text{minimize}} \ & \frac{1}{2}\sum_{u,i}(1+\gamma r_{ui})\big(p_{ui}-\mathbf{X}_u^T\mathbf{Y}_i\big)^2 \\ & +\lambda\big(\sum_u\|\mathbf{X}_u\|_2^2+\sum_i\|\mathbf{Y}_i\|_2^2\big). \end{aligned} \tag{1}$$

Vectors $\mathbf{X}_u, \mathbf{Y}_i \in \mathbb{R}^k$ are latent factors for user u and item i. User u's preference for i is determined by binarizing rating $r_{ui} \geq 0$

$$p_{ui} = \begin{cases} 1, & r_{ui} > 0, \\ 0, & r_{ui} = 0. \end{cases} \tag{2}$$

In this model, the lowest rating assigned to an item a user has observed is 1, so $r_{ui} = 0$ and $p_{ui} = 0$ if u has never observed item i. This model, accounts for all user-item pairs. Parameters γ, which scales the strength of user-item ratings, λ, which regularizes matrix factors, and k, the dimension of the latent space, are chosen by experiment. We extend this factorization to fit user-group engagement, solving

$$\begin{aligned} \underset{\mathbf{w},\mathbf{X}_u,\mathbf{Y}_g}{\text{minimize}} \ & \frac{1}{2}\sum_{u,g}(1+\gamma r_{ug})\big(p_{ug}-\mathbf{q}_{ug}^T\mathbf{w}-\mathbf{X}_u^T\mathbf{Y}_g\big)^2 \\ & +\alpha_u\sum_u\|\mathbf{X}_u\|_1+\frac{1}{2}\beta_u\sum_u\|\mathbf{X}_u\|_2^2 \\ & +\alpha_g\sum_g\|\mathbf{Y}_g\|_1+\frac{1}{2}\beta_g\sum_g\|\mathbf{Y}_g\|_2^2 \\ & +\alpha_w\|\mathbf{w}\|_1+\frac{1}{2}\beta_w\|\mathbf{w}\|_2^2 \\ \text{subject to } \ & \mathbf{X}_u,\mathbf{Y}_g \geq 0. \end{aligned} \tag{3}$$

Consistent with (1), $\mathbf{X}_u, \mathbf{Y}_g \in \mathbb{R}^k$ are user and group latent factors. We use either producing engagement or consuming engagement defined in Section as the implicit user-group rating r_{ug}, binarizing r_{ug} to user-group affinity p_{ug}. Parameters γ and k are rating sensitivity and latent factor dimension, and $\alpha_{\{u,g,w\}}, \beta_{\{u,g,w\}}$ are regularization parameters, all chosen by experimentation.

Variables $\mathbf{w} \in \mathbb{R}^3$ provide an optional bias for $\mathbf{q}_{ug} = [\bar{f}_u, \bar{c}_u, \bar{m}_g]^T$, measuring

- user properties:
  f_u: number of favorites user u has made
  c_u: number of comments user u has made
- group property:
  m_g: number of members in group g

These properties reflect overall levels of user and group activity. Note that we use normalized quantities

$$\bar{f}_u = \frac{f_u}{\sum_u f_u}, \quad \bar{c}_u = \frac{c_u}{\sum_u c_u}, \quad \bar{m}_g = \frac{m_g}{\sum_g m_g}, \tag{4}$$

to ensure that $\mathbf{q}_{ug}$ will be bounded in $(0, 1)$, reflecting the relative levels of user/group activity, and that elements of $\mathbf{w}$ are on the same scale. We refer to the special case where we fix $\mathbf{w} = 0$ as the unbiased model. Note that when we solve the unbiased model, we apply the optimizations reported in Hu's work (Hu, Koren, and Volinsky 2008), as well as the use of a tall and skinny QR-factorization (Constantine and Gleich 2011), to compute the matrix products of the form Y^TY which appear in the optimization procedure.

As our affinities $p_{ug} \geq 0$, we are motivated by previous work on non-negative matrix factorization (NMF) to compute $\mathbf{X}_u, \mathbf{Y}_g$ that are non-negative and sparse. As noted in the NMF literature – see (Lee and Seung 2001), for example – avoiding cancellation of factors of different signs, particularly if those factors are sparse, tends to produce a factorization that is more easily interpreted. Regarding each dimension of the user and group vectors as a topic, the suggestion that $\mathbf{X}_u, \mathbf{Y}_g$ should be non-zero in only a few coordinates reflects the notion that each user and group description is dominated by a few topical preferences.

As noted elsewhere (Hoyer 2004), the ratio between the l_1 and l_2 norms provides a measurement for vector sparsity. Rather than specify the sparsity explicity, we manage sparsity by regularizing against both the l_1 and l_2 norms, employing elastic-net regularization (Zou and Hastie 2005). This strategy reduces the model's sensitivity to the dimension of the latent factor dimension.

Having solved (3), we compute

$$\hat{p}_{ug} = [\bar{f}_u, \bar{c}_u, \bar{m}_g]^T\mathbf{w} + \mathbf{X}_u^T\mathbf{Y}_g, \tag{5}$$

as our prediction of user-group affinity. The largest predictions over user-group pairs for which $r_{ug} = 0$ are our recommendations.

Optimization Procedure

Problem (3) is convex in each of $\mathbf{X}_u, \mathbf{Y}_g$ and $\mathbf{w}$ separately, and so we use pathwise coordinate descent (Friedman et al. 2007) to optimize for each of these collections of variables iteratively. Recognizing that (3) is not jointly convex in all variables, we apply relaxation at each iteration, combining updated values with previous ones. This common strategy can increase the number of iterations, but prevents the algorithm from stalling at local minima.

Our algorithm follows a standard pattern for coordinate-wise optimization:

1. Initialize $\mathbf{Y}_g$, $\mathbf{w}$.
2. With $\mathbf{Y}_g$ and $\mathbf{w}$ fixed, solve (3) for $\mathbf{X}_u$.

3. With $\mathbf{X}_u$ and $\mathbf{w}$ fixed, solve (3) for $\mathbf{Y}_g$.
4. With $\mathbf{X}_u$ and $\mathbf{Y}_g$ fixed, solve (3) for $\mathbf{w}$.
5. If not converged, go to step 2; otherwise, stop.

Initialization Set $\mathbf{Y}_g = (1/k)[1, 1, \ldots, 1]^T$ and $\mathbf{w} = [0.1, 0.1, 0.1]^T$.

Minimization with respect to $\mathbf{X}_u$ Let $E(u)$ be the set of groups for which $r_{ug} > 0$. With $\mathbf{Y}_g, \mathbf{w}$ fixed, for each u, minimization (3) becomes

$$\begin{aligned} \underset{\mathbf{X}_u}{\text{minimize}} \;\; & \frac{1}{2}\mathbf{X}_u^T\Big(\sum_g \mathbf{Y}_g\mathbf{Y}_g^T + \sum_{g\in E(u)} \gamma r_{ug}\mathbf{Y}_g\mathbf{Y}_g^T\Big)\mathbf{X}_u \\ & - \Big(\sum_g (1+\gamma r_{ug})(p_{ug} - \mathbf{q}_{ug}^T\mathbf{w})\mathbf{Y}_g\Big)^T\mathbf{X}_u \\ & + \alpha_u\|\mathbf{X}_u\|_1 + \frac{1}{2}\beta_u\|\mathbf{X}_u\|_2^2 \\ \text{subject to} \;\; & \mathbf{X}_u \geq 0. \end{aligned} \tag{6}$$

The $\mathbf{X}_u$ can be computed in parallel, observing that $(p_{ug} - \mathbf{q}_{ug}^T\mathbf{w})$ and $\sum_g \mathbf{Y}_g\mathbf{Y}_g^T$ can be precomputed and re-used by each $\mathbf{X}_u$ calculation. A tall and skinny QR-factorization speeds up the $\sum_g \mathbf{Y}_g\mathbf{Y}_g^T$ calculation significantly.

The calculations for the $\mathbf{X}_u$ are sign-constrained elastic-net problems of the form

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \;\; & \frac{1}{2}\mathbf{x}^TA\mathbf{x} - \mathbf{b}^T\mathbf{x} + \alpha\|\mathbf{x}\|_1 + \frac{1}{2}\beta\|\mathbf{x}\|_2^2 \\ \text{subject to} \;\; & \mathbf{x} \geq 0, \end{aligned} \tag{7}$$

which can be solved by pathwise coordinate descent (Friedman et al. 2007).

Writing $\mathbf{X}_u^{(k)}$ as user latent factors at iteration k, and $\tilde{\mathbf{X}}_u^{(k+1)}$ latent factors computed by (6), we relax the update, setting our new latent factors as

$$\mathbf{X}_u^{(k+1)} = \mathbf{X}_u^{(k)} + \theta\left(\tilde{\mathbf{X}}_u^{(k+1)} - \mathbf{X}_u^{(k)}\right). \tag{8}$$

Where θ is the learning rate which controls the step-size in the parameter space. Choosing $\theta = 0.9$ appears to be a good compromise between fast convergence and local minima avoidance in relaxing updates for all of $\mathbf{X}_u, \mathbf{Y}_g$ and $\mathbf{w}$.

Minimization with respect to $\mathbf{Y}_g$ With the symmetry between $\mathbf{X}_u$ and $\mathbf{Y}_g$, minimization with respect to $\mathbf{Y}_g$ is the same as that of $\mathbf{X}_u$.

Minimization with respect to $\mathbf{w}$ With $\mathbf{X}_u, \mathbf{Y}_g$ fixed, optimization with respect to $\mathbf{w}$ becomes

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \;\; & \frac{1}{2}\mathbf{w}^T\Big(\sum_{ug}(1+\gamma r_{ug})\mathbf{q}_{ug}\mathbf{q}_{ug}^T\Big)\mathbf{w} \\ & - \Big(\sum_{ug}(1+\gamma r_{ug})(p_{ug} - \mathbf{X}_u^T\mathbf{Y}_g)\mathbf{q}_{ug}\Big)^T\mathbf{w} \\ & + \alpha_w\|\mathbf{w}\|_1 + \frac{1}{2}\beta_w\|\mathbf{w}\|_2^2 \\ \text{subject to} \;\; & \mathbf{w} \geq 0. \end{aligned} \tag{9}$$

This is also an elastic-net problem. Although the use of the bias term improves predicted user-group affinities, this expression shows that it comes at the cost of summing over all users and groups.

Temporal Model

To capture the time-varying nature of user preferences, group properties, and global biases, we collect group interaction data in discrete time intervals, forming time-dependent user-group engagement matrices. To estimate user-group affinity at time T, we solve problem (3) at times $T-1, T-2, \ldots, T-p$. We model the trajectories of user, group, and bias vectors using auto-regression (AR) and vector auto-regression (VAR) (Brockwell and Davis 2002; Hamilton 1994), extrapolating parameters and solutions to time T. Prediction (5) combines extrapolations to compute $\hat{p}_{ug}(T)$ and hence provide recommendations.

In our application of AR and VAR to our temporal model, we extend the user and group properties and factors to a time-varying equivalents:

- time-varying user properties:
  $f_u(t)$ is the number of favourites user u made in time interval t
  $c_u(t)$ is the number of comments user u made in time interval t
- time-varying group property:
  $m_g(t)$ is the number of new members in group g in time interval t
- time-varying factors:
  $\mathbf{X}_u(t)$ is user u's latent factors in time interval t
  $\mathbf{Y}_g(t)$ is group g's latent factors in time interval t

Our assumption is that both user preferences and group characteristics change gradually as time goes by, and there are trends for the changes in user-group engagement that we can use to predict future user-group affinity.

We divide the whole time span of data into T discrete time intervals (such as weeks, months or quarters). Users and groups will have their own qualities and latent factors in different time intervals. Note that the static model treats the whole time span as one single time interval, and can be regarded as a special case of the temporal model.

In particular, we adopt p-order auto-regression (AR(p)) to extrapolate the future qualities and p-order vector auto-regression (VAR(p)) to extrapolate the future latent factors. For time-varying functions $x(t)$, auto-regression assumes

$$x(T) = \sum_{k=1}^{p}\phi(k)x(T-k) + \epsilon(T), \tag{10}$$

where $\phi(k), k = 1 \ldots p$ are parameters we fit, and ϵ is the error in our time T estimate. Where $x(t)$ are scalar-valued, that is, where we fit $f_u(t), c_u(t)$ and $m_g(t)$ by AR(p), the $\phi(k)$ are scalar-valued; where $x(t)$ are vector-valued, that is, where we fit $\mathbf{X}_u, \mathbf{Y}_g$, and $\mathbf{w}$ by VAR(p), the $\phi(k)$ are matrix-valued.

We solve for parameters $\phi(k)$ by a least-squares minimization,

$$\underset{\phi(k)}{\text{minimize}} \sum_{t=p+1}^{T}\left[x(t) - \sum_{s=1}^{p}\phi(s)x(t-s)\right]^2. \tag{11}$$

This cost function represents a window of length p that passes forward over the time-varying data. We choose parameters $\phi(k)$ to best fit data at each time step $t \in p \ldots T-1$

based on the previous p time steps. Where (11) represents VAR(p), the squared term in square brackets is understood to represent the $l2$ vector norm.

Using data from $T-1$ trial intervals, we predict the user-group affinity at time T as

$$\hat{p}_{ug}(T) = [f_u(T), c_u(T), m_g(T)]^T \mathbf{w}(T) + \mathbf{X}_u(T)^T \mathbf{Y}_g(T) \tag{12}$$

In assuming that our parameter trajectories are smooth, such that user, group and global properties do not change abruptly, our future predictions directly leverage a long history of behaviour. This is in contrast to the Markovian assumption employed elsewhere (Zhang et al. 2014).

Experiments

Our experiments consider DeviantArt users and groups from 5 May 2011 to 31 August 2014. Given that the majority of DeviantArt users are young and still in school, we divide this 40-month time span into 10 equal 4-month intervals, each of which is roughly a school semester. The first 9 intervals serve as training/validation data; we withhold the last 4 months for testing. The static model aggregates the first 9 intervals into a single training-validation set, on which a 5-fold cross validation is used to select the optimal parameters, while the temporal model performs the matrix factorization on each of the first 9 intervals separately, making recommendations on the 10th interval using our temporal scheme.

User-Group Matrices

We examine two sets of engagement matrices, those for consumption engagement, geared towards art curators and viewers, and those for production engagement, geared towards artists. Production and consumption engagements are computed according to their definitions in the previous section.

Production engagement matrix: We filter the user-group matrix such that each user has joined at least one group and each group has at least one member by May 2011, and every user has at least one production engagement with some group in each 4-month interval. We omit groups that do not receive at least two production engagements in each interval. After performing this filtering, the production engagement matrix contains 8423 users and 4579 groups.

Consumption engagement matrix: Just as in production engagement, we filter the matrix to ensure that each user has joined at least one group and each group has at least one member by May 2011. We omit users having fewer than five consumption engagements in every time interval, and we omit groups having fewer than ten consumption engagements in every time interval. After filtering, the consumption engagement data contains 20328 users and 8772 groups.

Boolean membership matrix: For comparison, we also produce Boolean matrices corresponding to each of the filtered production and consumption matrices, setting

$$p_{ug} = \begin{cases} 1 & r_{ug} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

We set the confidence scaling parameter $\gamma = 0$ for experiments with Boolean membership matrices.

Table 1 summarizes the statistics of our data.

Table 1: Summary of production and consumption engagement matrices

	production	consumption
number of users	8423	20328
number of groups	4579	8772
number of non-zeros	161767	670662
matrix density	0.0042	0.0038

Recommendation Methods

To examine the contribution of each aspect of our model to the overall performance, we run several variants of our method:

- **Static**
  - **Unbiased:** Our static model with fixed $\mathbf{w} = 0$ in (3).
  - **Biased:** Our static model with variable bias $\mathbf{w}$
  - **Biased (L2 Norm):** Our static model with only $l2$-regularization but with variable bias $\mathbf{w}$
- **Temporal**
  - **Unbiased:** Our temporal model with fixed $\mathbf{w} = 0$ in (3)
  - **Biased:** Our temporal model with variable bias $\mathbf{w}$

We also include four baselines for comparison:

- **Popularity:** A baseline to indicate the problem's difficulty, recommending the K most popular groups; every user gets the same recommendations.
- **Boolean Membership:** The same matrix factorization method as Static Unbiased but applied to the Boolean membership matrices derived from the production and consumption engagement data.
- **Hu:** Uses implicit data according to Hu's method (2008), but in a static manner and without any of our improvements. Compared to Static Unbiased, this method omits the $l1$-regularization and the non-negativity constraint.
- **TMF:** We evaluate a version of Temporal Probabilistic Matrix Factorization (Zhang et al. 2014). As discussed before, TMF can be considered as a special case of our proposed method. Therefore, we implement the TMF concepts in our framework as follows. As in the original, we only model user factors' temporal evolution in this method, and restrict the calculation of a factor at T to take into account only the factor at time $T-1$. In contrast to the original work, we employ single time step vector autoregression rather than first order Markov chains. Lastly, we employ the method of Hu et al. (2008) to deal with implicit feedback.

As discussed in related work, other existing time-dependent methods only capture past temporal patterns rather than extrapolate future temporal dynamics to estimate future changes in users preferences, therefore we do not include them in our comparison baselines. Also, we don't compare our methods with Zeng's work (2013) because their work makes use of extra side information such as user-item rating

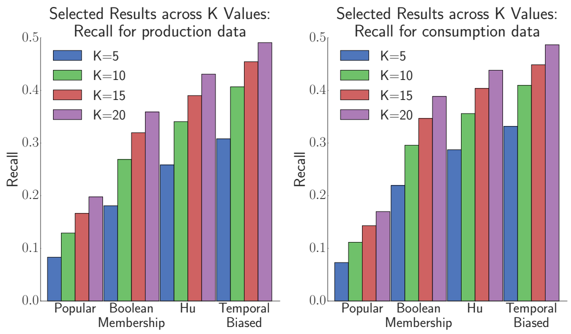
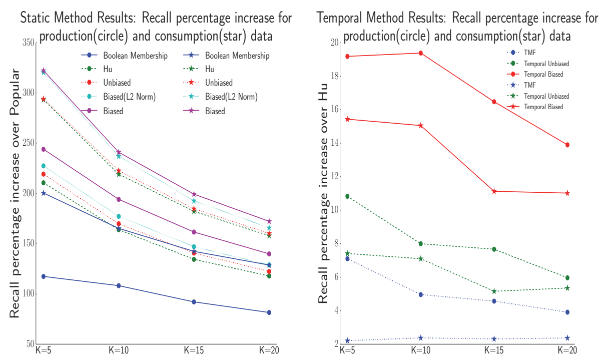


Figure 1: Recall across a range of K values for baselines (Popular, Boolean Membership and Hu) and our full proposed method (Temporal Biased)



(a) (b)

Figure 2: Recall for the static methods expressed as a percentage increase over the Popular baseline (a) and for the temporal methods expressed as a percentage increase over the Hu baseline (b)

and user-user friendship to improve recommendation accuracy. Our datasets don't contain rating or friendship information and their datasets don't have user-group engagement or user/group activity level. This being the case, their work is orthogonal to ours, so we are not able to compare our proposed methods with theirs either on our datasets or theirs.

Results

We use *Recall@K* and *Precision@K* in (Cremonesi, Koren, and Turrin 2010) to compare the performance of our methods with the baseline techniques. Grid search is used to explore various values of p in $[1, 8]$ and we present results with $p = 3$ which achieve the best performance. Figure 1 highlights the general performance of our full Temporal Biased method relative to some of the baseline comparison methods, demonstrating that the method drastically outperforms a simple Boolean Membership encoding, and also improves substantially upon a standard static implicit approach (Hu) run on our engagement data. Figure 2(a) displays the results of all static methods as a percentage increase over the popular baseline. Clearly, using the more sophisticated implicit measurements of user engagement with groups pro-

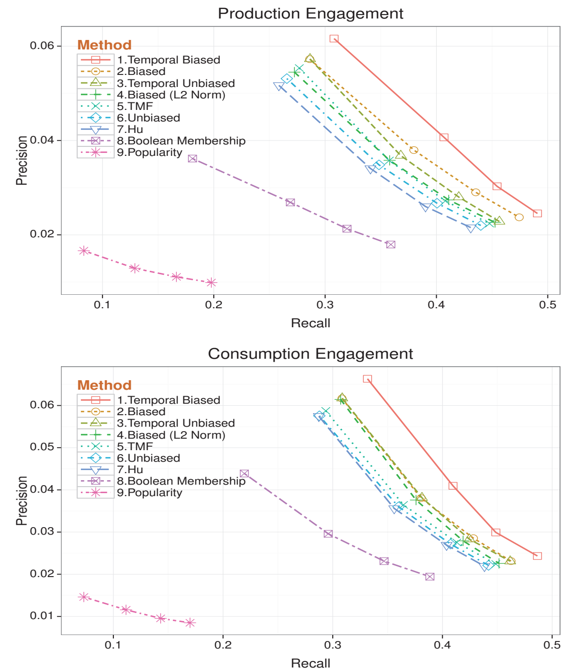


Figure 3: Results of precision versus recall for all methods, showing improved performance as complexity is added to the model

vides a substantial gain in recommendation performance. The minor changes in the form of non-negative elastic net regularization between Hu and our Static Unbiased method yield a slight improvement on the production data. Adding the biases leads to a substantial gain in recall performance, especially together with l_1-regularization in the production case. While using a bias also improves performance on the consumption engagement data, there is almost no gain compared to using only l_2 regularization with biases in this case. This difference might be explained by the tendency of artists to have focused areas of interest and ability, and hence be well-modelled by sparse latent vectors. Conversely, art curators need not be so focused in their range of tastes, and hence may be better modelled by a less sparse latent preference factorization. Figure 2(b) summarizes the results of the temporal baselines and methods as percentage improvements over the implicit engagement data based static Hu baseline. While TMF shows an improvement from taking the last time slice into account for user factors, modelling all the past time slices through smooth functions lets our temporal methods clearly outperform this baseline. Figure 3 presents the trade-off between recall and precision. Each line in Figure 3 reports the precision of a method at a given recall. Recall will usually go up as K increases while precision tends to go down, which confirms that a trade-off between recall and precision is unavoidable in top-K recommendations. Additionally, Figure 3 demonstrates that the Temporal Biased method also obtains the best performance in terms of precision, followed by other variants of our method. We observe that the relative performances of all methods in terms of recall and precision are consistent on both datasets.

Conclusion

In this paper, we propose production engagement and consumption engagement as measures that are more fine-grained than Boolean user-group membership to quantify user-group affinity more accurately. We present a time-dependent matrix factorization model to recommend groups of users, and perform experiments on real-world user-group datasets from DeviantArt to demonstrate the improvement of our proposed method. The experimental results show that for the complex problem of recommending social network groups to users taking into account detailed implicit engagement data rather than simply Boolean group memberships yields substantial improvements in recommendation performance. We achieve another performance boost by modelling the evolving nature of the relationships between users and groups smoothly over time rather than assuming they are static or can be predicted from the last time slice alone.

References

Ali, I., and Kim, S.-W. 2015. Group recommendations: approaches and evaluation. In *Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication*, 105. ACM.

Baltrunas, L.; Makcinskas, T.; and Ricci, F. 2010. Group recommendations with rank aggregation and collaborative filtering. In *Proceedings of the fourth ACM conference on Recommender systems*, 119–126. ACM.

Brockwell, P. J., and Davis, R. A. 2002. *Introduction to Time Series and Forecasting*, volume 1. Taylor & Francis.

Chen, W.-Y.; Chu, J.-C.; Luan, J.; Bai, H.; Wang, Y.; and Chang, E. Y. 2009. Collaborative filtering for orkut communities: discovery of user latent behavior. In *Proceedings of the 18th International Conference on World Wide Web*, 681–690. ACM.

Chen, W.-Y.; Zhang, D.; and Chang, E. Y. 2008. Combinational collaborative filtering for personalized community recommendation. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data mining*, 115–123. ACM.

Chua, F. C. T.; Oentaryo, R. J.; and Lim, E.-P. 2013. Modeling temporal adoptions using dynamic matrix factorization. In *Data Mining (ICDM), 2013 IEEE 13th International Conference on*, 91–100. IEEE.

Constantine, P. G., and Gleich, D. F. 2011. Tall and skinny QR factorizations in MapReduce architectures. In *Proceedings of the Second International Workshop on MapReduce and its Applications*, 43–50. ACM.

Cremonesi, P.; Koren, Y.; and Turrin, R. 2010. Performance of recommender algorithms on top-n recommendation tasks. In *Proceedings of the Fourth ACM Conference on Recommender Systems*, 39–46. ACM.

DeviantArt. http://www.deviantart.com/.

Douban. http://www.douban.com/.

Friedman, J.; Hastie, T.; Höfling, H.; Tibshirani, R.; et al. 2007. Pathwise coordinate optimization. *The Annals of Applied Statistics* 1(2):302–332.

Hamilton, J. D. 1994. *Time series analysis*, volume 2. Princeton University Press.

Hoyer, P. O. 2004. Non-negative matrix factorization with sparseness constraints. *The Journal of Machine Learning Research* 5:1457–1469.

Hu, Y.; Koren, Y.; and Volinsky, C. 2008. Collaborative filtering for implicit feedback datasets. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, 263–272. IEEE.

Koenigstein, N.; Dror, G.; and Koren, Y. 2011. Yahoo! music recommendations: modeling music ratings with temporal dynamics and item taxonomy. In *Proceedings of the fifth ACM conference on Recommender systems*, 165–172. ACM.

Koren, Y. 2010. Collaborative filtering with temporal dynamics. *Communications of the ACM* 53(4):89–97.

Lathia, N.; Hailes, S.; and Capra, L. 2009. Temporal collaborative filtering with adaptive neighbourhoods. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, 796–797. ACM.

Lee, D. D., and Seung, H. S. 2001. Algorithms for non-negative matrix factorization. In *Advances in Neural Information Processing Systems*, 556–562.

Salakhutdinov, R., and Mnih, A. 2008a. Bayesian probabilistic matrix factorization using Markov chain Monte Carlo. In *Proceedings of the 25th International Conference on Machine Learning*, 880–887. ACM.

Salakhutdinov, R., and Mnih, A. 2008b. Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems*, volume 20.

Vasuki, V.; Natarajan, N.; Lu, Z.; and Dhillon, I. S. 2010. Affiliation recommendation using auxiliary networks. In *Proceedings of the fourth ACM conference on Recommender Systems*, 103–110. ACM.

Wang, J.; Zhao, Z.; Zhou, J.; Wang, H.; Cui, B.; and Qi, G. 2012. Recommending flickr groups with social topic model. *Information retrieval* 15(3-4):278–295.

Xiang, L.; Yuan, Q.; Zhao, S.; Chen, L.; Zhang, X.; Yang, Q.; and Sun, J. 2010. Temporal recommendation on graphs via long-and short-term preference fusion. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 723–732. ACM.

Xiong, L.; Chen, X.; Huang, T.-K.; Schneider, J. G.; and Carbonell, J. G. 2010. Temporal collaborative filtering with Bayesian probabilistic tensor factorization. In *SDM*, volume 10, 211–222. SIAM.

Zeng, W., and Chen, L. 2013. Recommending interest groups to social media users by incorporating heterogeneous resources. In *Recent Trends in Applied Artificial Intelligence*. Springer. 361–371.

Zhang, C.; Wang, K.; Yu, H.; Sun, J.; and Lim, E.-P. 2014. Latent factor transition for dynamic collaborative filtering. In *Proceedings of the SIAM International Conference on Data Mining*. SIAM.

Zou, H., and Hastie, T. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67(2):301–320.