When Personalization Meets Conformity: Collective Similarity based Multi-Domain Recommendation

Xi Zhang¹, Jian Cheng¹, Shuang Qiu¹, Zhenfeng Zhu², Hanqing Lu¹ ¹National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences ²Institute of information Science, Beijing Jiaotong University {xi.zhang, jcheng, shuang.qiu, luhq}@nlpr.ia.ac.cn, zhfzhu@bjtu.edu.cn

ABSTRACT

Existing recommender systems place emphasis on personalization to achieve promising accuracy. However, in the context of multiple domain, users are likely to seek the same behaviors as domain authorities. This conformity effect provides a wealth of prior knowledge when it comes to multidomain recommendation, but has not been fully exploited. In particular, users whose behaviors are significant similar with the public tastes can be viewed as domain authorities. To detect these users meanwhile embed conformity into recommendation, a domain-specific similarity matrix is intuitively employed. Therefore, a collective similarity is obtained to leverage the conformity with personalization. In this paper, we establish a Collective Structure Sparse Representation(CSSR) method for multi-domain recommendation. Based on adaptive k-Nearest-Neighbor framework, we impose the lasso and group lasso penalties as well as least square loss to jointly optimize the collective similarity. Experimental results on real-world data confirm the effectiveness of the proposed method.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Information Filtering

Keywords

Multiple Domains; Recommendation; Conformity

1. INTRODUCTION

The fast growth of Web 2.0 technologies facilitates and encourages various online user behaviors, meanwhile brings tremendous information to the public. Personalized recommendation, as critical methods to push the right information to the right users, have attracted large amounts of research in both industry and academia. Among these methods, Collaborative Filtering(CF) has superior performance than other methods via an underlying assumption about entity sim-

SIGIR'15, August 09 - 13, 2015, Santiago, Chile,

© 2015 ACM. ISBN 978-1-4503-3621-5/15/08 ...\$15.00.

DOI: http://dx.doi.org/10.1145/2766462.2767810.

ilarity. Namely, users with common interests in the past would behave much more similarly on items in the future. Traditionally, CF approaches regard mining distinct types of user behavior as a task within a single domain. Nevertheless, in many real applications, multiple behaviors ranging from reading books, favoring music, to watching videos can reflect the characteristics of users together, and thus training a group of predictors in a unified manner would enhance accuracy for all the domains.

To cope with multi-domain recommendation, a straightforward scheme is to extend the CF algorithms into a multitask learning problem solved by propagating similarities among domains [4, 2]. However, somewhat surprisingly, the context of multi-domain often introduces *conformity*¹ to online user behaviors, which cannot be ignored. In particular, the public are inclined to conform to the choices or beliefs of domain authorities [6]. For example, comments from well-known movie critics would be likely to affect the trend of box offices while reading lists of celebrated writers would probably receive admiration. That is to say, compared with ordinary users, opinions of these domain elites are more representative and thus have considerable influence on recommendation results. Unfortunately, this effect has not been fully exploited in existing multi-domain approaches.

The key problem becomes how to detect the domain authorities and integrate the conformity effect into recommendation. To simplify this problem, we concern that the elites are users whose past behaviors can largely reflect the public tastes. In fact, if a user has significant similar behaviors to a large number of other users, she/he is probably the elite user who can represent others. To this end, a domainspecific user similarity matrix based on observed behaviors is built to embody conformity. On the other hand, since users' intrinsic preferences over domains are also decisive, a domain-shared user similarity matrix is utilized to profile users globally. Therefore, a collective similarity combining the effect of conformity with the original personalization is established.

In this paper, we propose a Collective Structured Sparse Representation (CSSR) method to optimize the collective similarity for a top-N multi-domain recommendation task. The CSSR is on the basis of the adaptive k-Nearest-Neighbor framework. To learn ranked item lists under each domain for users, a least square regression loss of feedback representation is employed. To encode personalization and confor-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

¹The general conformity is a social phenomenon of matching attitudes, beliefs, and behaviors to group norm. Here we narrow the concept to achieving same behaviors as elites.

mity, regularization with lasso and group lasso are adopted. Also, we show how to use the Alternating Direction Method of Multipliers(ADMM) [1] to efficiently optimize the parameters. Experiments show that our method consistently achieves more accurate results than other state-of-the-arts.

2. THE PROPOSED APPROACH

2.1 Notation and Background

Suppose that there are a set of users $\mathcal{U} = \{u_1, \dots, u_n\}$ and D types of domains. Let the matrix $\mathbf{X}_d \in \mathbb{R}^{m_d \times n}$ record the interactions between the overall user set and item set \mathcal{V}_d for the d-th domain, where $d = 1, \dots, D$ and m_d is the size of the item set. Then if a user has any behavior on item, the element entry of \mathbf{X}_d could be 1 or a positive value, otherwise the entry is set as 0. Note that implicit feedback such as clicks, purchases and posts are easier to obtain than rating records in real-world scenarios. Thus we assume that the multiple feedback matrices are binary.

For a multiple domain recommendation, given a series of user-item feedback matrices $\{\mathbf{X}_1, \cdots, \mathbf{X}_D\}$, our goal is to recommend a personalized ranking list of the potential enjoyed items under each domain for users.

Here we introduce a general similarity based framework for recommendation without multi-domain. Normally, preference scores are estimated by a function of $f(\mathbf{X}, \Theta)$, where **X** is the feedback matrix and Θ is the model parameter. The adaptive k-Nearest-Neighbor(kNN) method, which is very popular for collaborative filtering, has been shown its simpleness but effectiveness in top-N recommendation [3]. We establish our model based on the adaptive kNN technique in this paper. Consider a user based kNN problem, the parameter Θ is set as matrix $\mathbf{W} \in \mathbb{R}^{n \times n}$, and the prediction function $f(\cdot)$ could be presented as an aggregation of the feedback values of k nearest neighbors. Mathematically, $\hat{x}_{ij} = \mathbf{x}_i^T \mathbf{w}_j$, where \hat{x}_{ij} is the predicted scores, $\mathbf{x}_i \in \mathbb{R}^n$ denotes the feedback indicator of item v_i , and $\mathbf{w}_i \in \mathbb{R}^n$ denotes a similarity coefficients of user u_i . The parameter **W** can be regarded as the similarity between user pairs, which can be optimized by the following problem

$$\widetilde{\mathbf{W}} = \arg\min_{\mathbf{W}} \mathcal{L}(\mathbf{W}) + \Omega(\mathbf{W}) \tag{1}$$

where \mathcal{L} is a loss function usually defined as least squares or logistic regression, and Ω is a regularization penalty to enforce parameter with specific structures. Ridge, lasso, or elastic net regularization has been used in previous methods.

2.2 Collective Similarity

Now we consider the problem of multi-domain recommendation. As the user set is common across domains, the user based similarity is exploited. The interpretation behind the idea is simple. That is, to learn a set of nearest neighbors who are sufficiently similar with the user in terms of multiple behaviors over domains. In other words, a user's interests can be represented by the set of her/his neighbors. The usual approach assumes that all users are ordinary with the comparative ability of representation and thus the personalized neighbors are mined for each users independently. Nevertheless, conformity always occurs in the context of multi-domain. Under a specific domain, elite users' favorites are more acceptable to the public during recommendation. This indicates that the similarity coefficients between these



Figure 1: Illustration of the CSSR Model

representative elite users and others tend to be significant. Hence, we propose a collective similarity as

$$\mathbf{W}_d = \mathbf{V} + \mathbf{V}_d \tag{2}$$

where $d = 1, \dots, D$. Parameter matrix $\mathbf{V} \in \mathbb{R}^{n \times n}$ indicates the personalized neighbors with intrinsic similar interest across domains, which is a domain-shared component. Parameter matrix $\mathbf{V}_d \in \mathbb{R}^{n \times n}$ embodies the representative of elite neighbors in the domain, which is a domain-specific component.

Next, we seek to explore the conformity by regularizing the structure of the parameter set $\{\mathbf{V}_1, \dots, \mathbf{V}_D\}$ with group lasso. On the one hand, the domain-specific elite users can largely represent other users, which results in more dense for the corresponding rows in \mathbf{V}_d . On the other hand, the fact that elite users are always the minority leads to the sparsity in columns. Hence, the above structure serves to define a group lasso regularization term

$$\Omega_{glas}(\mathbf{V}_d) = \beta_d \|\mathbf{V}_d\|_{2,1} = \sum_{i=1}^n \beta_d \|\mathbf{v}_d^i\|_2 \tag{3}$$

where $\|\mathbf{v}_d^i\|_2$ is the ℓ_2 -norm for each row of \mathbf{V}_d . $\ell_{2,1}$ -norm encourages the row sparsity with jointly selecting the significant rows as elite users. β_d encodes the contributions of group lasso for each domain. Without loss of generality, we assume $\forall d, \beta_d = \beta$.

Furthermore, users' personalized neighbors are also sparse in the overall user set. A lasso constraint is utilized to filter out noised coefficients in \mathbf{V} , which is

$$\Omega_{las}(\mathbf{V}) = \lambda \|\mathbf{V}\|_1 = \sum_{i=1}^n \sum_{j=1}^n \lambda |v_{ij}|$$
(4)

With above discussion, the collective similarity can be incorporated into a loss function with a least square loss as $\mathcal{L}(\mathbf{W}_d) = \frac{1}{2} \|\mathbf{X}_d - \mathbf{X}_d \mathbf{W}_d\|_F^2$. Following the general framework in Eq. (1), we propose a CSSR model for adaptive kNN problem in multiple domains as

$$\min_{\Theta} \sum_{d=1}^{D} \alpha_d \left(\mathcal{L}(\mathbf{W}_d) + \Omega_{glas}(\mathbf{V}_d) \right) + \Omega_{las}(\mathbf{V})
s.t. \mathbf{W}_d = \mathbf{V} + \mathbf{V}_d, \mathbf{W}_d \ge 0, \operatorname{diag}(\mathbf{W}_d) = 0
\forall d, \quad d = 1, \cdots, D$$
(5)

The parameter set Θ is written as $\{\mathbf{W}_d, \mathbf{V}_d, \mathbf{V}\}$ for succinct. And α_d balances the effect of different domain. In addition, the non-negative constraint is used on similarity for a more interpretable consideration, while $\mathbf{diag}(\mathbf{W}_d) = 0$ ensures that the trivial solution of identity matrix is avoided. From the unified objective function, information of feed-

back matrices can be transferred among domains by \mathbf{V} and distinctions can be captured by \mathbf{V}_d . The entire model is depicted in Figure 1.

2.3 Optimization

Although non-smooth terms are introduced, our approach maintains convexity of the objective function. We propose to apply the ADMM [1] to split the original problem into several subproblems which can be handled in alternating directions.

To solve the problem in Eq. (5), we first obtain the augmented Lagrangian as

$$\min_{\Theta, \mathbf{Y}_{d}} \sum_{d=1}^{D} \alpha_{d} \left(\mathcal{L}(\mathbf{W}_{d}) + \mathcal{I}_{+}(\mathbf{W}_{d}) + \Omega_{glas}(\mathbf{V}_{d}) \right) + \Omega_{las}(\mathbf{V}) + \sum_{d=1}^{D} \left(tr(\mathbf{Y}_{d}^{T}(\mathbf{W}_{d} - \mathbf{V} - \mathbf{V}_{d})) + \frac{\rho}{2} \|\mathbf{W}_{d} - \mathbf{V} - \mathbf{V}_{d}\|_{F}^{2} \right)$$
(6)

where \mathbf{Y}_d is Lagrangian multiplier of the *d*-th domain, $\mathcal{I}_+(\cdot)$ is the indicator function for the non-negative constraint, and $\rho > 0$ is a penalty parameter. For the given domain *d*, optimizing { $\mathbf{W}_d, \mathbf{V}_d, \mathbf{Y}_d$ } are independent with optimizing the set of variables in other domains, thus ADMM proceeds by solving following problems alternately until convergence

$$\min_{\mathbf{W}_d \ge 0} \alpha_d \mathcal{L}(\mathbf{W}_d) + \frac{\rho}{2} \|\mathbf{W}_d - \mathbf{V} - \mathbf{V}_d + \mathbf{U}_d\|_F^2$$
(7a)

$$\min_{\mathbf{V}_d} \alpha_d \Omega_{glas}(\mathbf{V}_d) + \frac{\rho}{2} \|\mathbf{W}_d - \mathbf{V} - \mathbf{V}_d + \mathbf{U}_d\|_F^2 \qquad (7b)$$

$$\min_{\mathbf{V}} \Omega_{las}(\mathbf{V}) + \frac{\rho}{2} \sum_{d=1}^{D} \|\mathbf{W}_d - \mathbf{V} - \mathbf{V}_d + \mathbf{U}_d\|_F^2 \qquad (7c)$$

$$\mathbf{U}_d = \mathbf{U}_d + \mathbf{W}_d - \mathbf{V} - \mathbf{V}_d \tag{7d}$$

where $\mathbf{U}_d = \frac{1}{\rho} \mathbf{Y}_d$ leads to a scaled form. Then we provide closed-form solutions of the subproblems for Eq.(7a), Eq.(7b) and Eq.(7c) as below.

Update for \mathbf{W}_d. We first rewrite the problem into an unconstrained problem as

$$\min_{\mathbf{W}_d} \alpha_d \mathcal{L}(\mathbf{W}_d) + \frac{\rho}{2} \|\mathbf{W}_d - \mathbf{V} - \mathbf{V}_d + \mathbf{U}_d\|_F^2 - tr(\Phi_d^T \mathbf{W}_d)$$
(8)

Since Eq.(7a) is smooth, the solution can be found by taking its derivative and set it to be zero

$$\left(\mathbf{X}_{d}^{T}\mathbf{X}_{d}+\rho\mathbf{I}\right)\mathbf{W}_{d}+\rho(\mathbf{U}_{d}-\mathbf{V}-\mathbf{V}_{d})-\mathbf{X}_{d}^{T}\mathbf{X}_{d}-\Phi_{d}=0$$
 (9)

Using the Karush-Kuhn-Tucker complementary condition for the nonnegativity of \mathbf{W}_d , $[\Phi_d]_{ij}[\mathbf{W}_d]_{ij} = 0$, we get

$$[\mathbf{W}_d]_{ij} = [\mathbf{W}_d]_{ij} \sqrt{\frac{\left[\mathbf{X}_d^T \mathbf{X}_d + \rho(\mathbf{U}_d - \mathbf{V} - \mathbf{V}_d)^-\right]_{ij}}{\left[(\mathbf{X}_d^T \mathbf{X}_d + \rho\mathbf{I})\mathbf{W}_d + \rho(\mathbf{U}_d - \mathbf{V} - \mathbf{V}_d)^+\right]_{ij}}}$$
(10)

To guarantee the nonnegativity, we decompose matrices with any signs as $\mathbf{A} = \mathbf{A}^+ - \mathbf{A}^-$, where $\mathbf{A}_{ij}^+ = (|\mathbf{A}_{ij}| + \mathbf{A}_{ij})/2$ and $\mathbf{A}_{ij}^- = (|\mathbf{A}_{ij}| - \mathbf{A}_{ij})/2$.

Update for \mathbf{V}_d . For each row of \mathbf{V}_d , the optimization problem in Eq.(7b) comes to

$$\min_{\mathbf{v}_d^i} \beta \|\mathbf{v}_d^i\|_2 + \frac{\rho}{2} \|\mathbf{v}_d^i - (\mathbf{w}_d^i - \mathbf{v}^i + \mathbf{u}_d^i)\|_2^2$$
(11)

where *i* is the row index of parameter matrices. Let $\mathbf{z}_{d}^{i} = \mathbf{w}_{d}^{i} - \mathbf{v}^{i} + \mathbf{u}_{d}^{i}$. This problem can be solved by a proximal operator of group lasso to each row vector as

$$\mathbf{v}_{d}^{i} = S_{glas,\frac{\beta}{\rho}}(\mathbf{z}_{d}^{i}) = \begin{cases} 0, & \text{if } \|\mathbf{z}_{d}^{i}\|_{2} \le \frac{\beta}{\rho} \\ \frac{\|\mathbf{z}_{d}^{i}\|_{2} - \frac{\beta}{\rho}}{\|\mathbf{z}_{d}^{i}\|_{2}} \mathbf{z}_{d}^{i}, & \text{otherwise.} \end{cases}$$
(12)

Update for V. Optimizing the second term with respect to **V** in Eq.(7c) equals to optimizing $\frac{\rho}{2} \|\mathbf{V} - \frac{1}{D} \sum_{d=1}^{D} (\mathbf{W}_d - \mathbf{V}_d + \mathbf{U}_d)\|_F^2$ which can be substituted into Eq.(7c). Then we have the element-wise optimal solution as

$$[\mathbf{V}]_{ij} = S_{\text{las},\frac{\lambda}{\rho}} \left(\left[\frac{1}{D} \sum_{d=1}^{D} (\mathbf{W}_d - \mathbf{V}_d + \mathbf{U}_d) \right]_{ij} \right)$$
(13)

where $S_{\text{las},\varepsilon}(\cdot)$ is the soft-thresholding operator

$$S_{\text{las},\varepsilon}(x) = \begin{cases} x - \varepsilon, & \text{if } x > \varepsilon \\ x + \varepsilon, & \text{if } x < -\varepsilon \\ 0, & \text{otherwise.} \end{cases}$$
(14)

3. EXPERIMENTS

3.1 Experiment Settings

To empirically study the effectiveness of our method, we perform experiments on a multi-domain dataset crawled from the publicly available site Douban². It is a famous Web2.0 website containing rating behavior of users, scaled from 1 to 5 stars, on books, music and movies. As we face with a top-N recommendation task, we use the implicit feedback instead of rating scores. For a sufficient evaluation of each user, we filtering out users with less than 10 feedbacks on the three domains and obtained a dataset of 5,916 users. The detailed description is presented in Table 1.

 Table 1: Description of Douban Dataset

| Domain | #Items | %Sparsity | #Ratings per User |
|--------|--------|-----------|-------------------|
| Book | 14,155 | 99.85 | 22 |
| Music | 15,492 | 99.75 | 38 |
| Movie | 7,845 | 98.87 | 88 |

For personalized item recommendation, we analyze performance of the model by comparing the top suggestions to the true behaviors taken by a user. We adopt two widely used evaluation metrics in top-N recommendation: MAP(Mean Average Precision), NDCG(Normalized Discounted Cumulative Gain) with the setting N = 5. Higher values on the metrics imply better recommendation results.

We compare the proposed CSSR with several popular baseline methods: PopRank, NCDCF_U, NCDCF_I [5], mrBPR [2], mrSLIM. Here we extend SLIM [3] to mrSLIM by constructing a feedback matrix that takes all items in three domains as rows. For a comprehensive comparison, performances of CSSRsh and CSSRsp are shown. They merely consider the domain-shared or domain-specific similarity in Eq. (2), and utilize lasso constrains.

In our experiments, we randomly pick 80% observed feedbacks for each user to form the training set in each domain and the rest of 20% is test set. The random sampling is repeated 10 times and the average performance are reported. We set the weights combination of domains in CSSR as

²http://www.douban.com

Table 2: Prediction performance (mean \pm std.) of three variants of CSSR and PopRank, NCDCF_U, NCD-CF_I, mrSLIM, mrBPR on three domains of Douban dataset. Results in bold indicate the best ones.

| Mothode | Params | | Book | | Music | | Movie | |
|-------------------------|--------|------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| Methods | | | MAP | NDCG | MAP | NDCG | MAP | NDCG |
| PopRank | | N/A | 0.1370 ± 0.0039 | $0.0619 {\scriptstyle \pm 0.0014}$ | 0.1879 ± 0.0029 | $0.0922 {\scriptstyle \pm 0.0014}$ | $0.3784 {\scriptstyle \pm 0.0031}$ | 0.2249 ± 0.0017 |
| NCDCF_U | - | 100 | 0.1826 ± 0.0033 | $0.0813 {\scriptstyle \pm 0.0013}$ | 0.2647 ± 0.0035 | $0.1325 {\scriptstyle \pm 0.0015}$ | 0.4982 ± 0.0038 | $0.3159 {\scriptstyle \pm 0.0024}$ |
| NCDCF_I | - | 50 | 0.1864 ± 0.0029 | $0.0835 {\scriptstyle \pm 0.0011}$ | 0.2668 ± 0.0016 | 0.1341 ± 0.0007 | $0.5049 {\scriptstyle \pm 0.0047}$ | $0.3230 {\scriptstyle \pm 0.0027}$ |
| mrSLIM | 5 | 0.1 | 0.2014 ± 0.0035 | $0.0906 {\scriptstyle \pm 0.0013}$ | 0.2861 ± 0.0028 | $0.1453 {\scriptstyle \pm 0.0022}$ | $0.5115 {\scriptstyle \pm 0.0040}$ | $0.3241 {\scriptstyle \pm 0.0023}$ |
| mrBPR | 0.01 | 1e-3 | $0.2333 {\scriptstyle \pm 0.0035}$ | 0.1071 ± 0.0019 | $0.3180 {\scriptstyle \pm 0.0037}$ | $0.1554 {\scriptstyle \pm 0.0021}$ | $0.5206 \scriptstyle \pm 0.0039$ | $0.3377 \scriptstyle \pm 0.0025$ |
| CSSRsh | - | 0.01 | 0.2217 ± 0.0023 | 0.1013 ± 0.0010 | 0.2949 ± 0.0028 | 0.1453 ± 0.0011 | 0.4952 ± 0.0055 | $0.3106 \scriptstyle \pm 0.0026$ |
| \mathbf{CSSRsp} | - | 0.01 | 0.2348 ± 0.0041 | 0.1070 ± 0.0019 | 0.3020 ± 0.0046 | $0.1539 {\scriptstyle \pm 0.0022}$ | $0.5052 {\scriptstyle \pm 0.0029}$ | $0.3241 {\scriptstyle \pm 0.0020}$ |
| CSSR | 0.5 | 0.01 | 0.2746 ± 0.0032 | $0.1241 \scriptstyle \pm 0.0013$ | $0.3345{\scriptstyle\pm0.0037}$ | $0.1703 \scriptstyle \pm 0.0014$ | 0.5341 ± 0.0030 | 0.3471 ± 0.0015 |



Figure 2: Performance variation of CSSR with respect to β and λ on three domains.

 $\{\alpha_1 = 0.2, \alpha_2 = 0.3, \alpha_3 = 0.5\}$ based on the MAP performance on the test set. The impact of group lasso parameter β and lasso parameter λ will be further studied.

3.2 Results and Analysis

Experimental results of above baselines and three variants of the proposed CSSR are shown in Table 2. The optimal parameters we obtained for CSSR are $\beta = 0.5, \lambda = 0.01$. From Table 2 we observed that, our full model CSSR consistently outperforms all the other baselines. Specifically, for the reason that PopRank is not a personalized algorithm and only recommends items based on their popularity, all other baselines can beat it. This proves the importance of personalization in recommender systems. By comparing the results along with three domains, we discover that recommending movies is a relatively easier task as all the approaches except PopRank perform well on this domain. Thereby, transferring knowledge from movie domain would benefit the other two tasks. Lacking of a properly designed mechanism for multi-task learning, NCDCF_U and NCDCF_I (with the number of nearest neighbors k = 100 and k = 50) cannot achieve satisfying results in book and music domains. We implement mrSLIM by setting the weights of elastic net as $\beta = 5$ and $\lambda = 0.1$ to find an optimal performance. With the learned item similarity matrix, mrSLIM behaves better than its simple version NCDCF_I. Moveover, CSSR surpasses the state-of-the-art ranking method mrBPR with the learning rate as 0.01 and the weights of ℓ_2 -norm as 1e - 3. mrBPR focuses on studying the domain consistency to model a sharing component, but ignores the heterogeneous part which might be different in domains, in particular, conformity.

To evaluate the effectiveness of collective similarity, we compare the three variants of CSSR. CSSRsh and CSSRsp encode consistency and heterogeneous among domains under the adaptive kNN settings, separately. However, the results show that they obtain improvement on Book and Music domains but fail to outperform previous methods in Movie domain. Integrating different aspect of the two variants, our method yields the best performance in multiple domains.

Finally, to understand the influence of regularization terms, we analyze the performance variations with respect to λ and β . As illustrated in Fig.2, by fixing α , the performance first increases when β gets larger. This verifies our conformity assumption that some users are representative and could be detected by enforcing a group lasso structure. But overwhelmingly large β would cause information loss in domains and result in the accuracy declining. By fixing β , we can see that the α is less sensitive. Another interesting observation is when β is improperly large($\beta = 10$), decreasing α can upgrade results to some extent. This indicates some missing information of domain-specific component could be compensated by domain-shared component.

4. CONCLUSIONS

In this paper, we propose a novel CSSR method for multidomain recommendation, which integrates personalization with conformity effect to construct a collective similarity parameter. By applying the least square loss, ranking scores can be predicted by the optimized neighbors. To model the different kind of neighbors, lasso and group lasso constraints are used. Experiments on multi-domain dataset show our method outperforms baselines for top-N recommendation.

5. REFERENCES

- S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 2011.
- [2] A. Krohn-Grimberghe, L. Drumond, C. Freudenthaler, and L. Schmidt-Thieme. Multi-relational matrix factorization using bayesian personalized ranking for social network data. In WSDM, 2012.
- [3] X. Ning and G. Karypis. SLIM: Sparse linear models for top-n recommender systems. In *ICDM*, 2011.
- [4] A. P. Singh and G. J. Gordon. Relational learning via collective matrix factorization. In *KDD*, 2008.
- [5] T. Yuan, J. Cheng, X. Zhang, S. Qiu, and H. Lu. Recommendation by mining multiple user behaviors with group sparsity. In AAAI, 2014.
- [6] X. Zhang, J. Cheng, T. Yuan, B. Niu, and H. Lu. Toprec: domain-specific recommendation through community topic mining in social network. 2013.