PRemiSE: Personalized News Recommendation via Implicit Social Experts

Chen Lin
Xiamen University
Xiamen, China
chenlin@xmu.edu.cn

Runquan Xie
Xiamen University
Xiamen, China
RunquanXie@stu.xmu.edu.cn

Lei Li
Florida International University
Miami, FL, U.S.A.
lli003@cs.fiu.edu

Zhenhua Huang
Tongji University
Shanghai, China
jukie.huang@gmail.com

Tao Li
Florida International University
Miami, FL, U.S.A.
taoli@cs.fiu.edu

ABSTRACT

A variety of news recommender systems based on different strategies have been proposed to provide news personalization services for online news readers. However, little research work has been reported on utilizing the implicit "social" factors (i.e., the potential influential experts in news reading community) among news readers to facilitate news personalization. In this paper, we investigate the feasibility of integrating content-based methods, collaborative filtering and information diffusion models by employing probabilistic matrix factorization techniques. We propose **PRemiSE**, a novel *P*ersonalized news *Re*commendation framework via i*m*pl*i*cit *S*ocial *E*xperts, in which the opinions of potential influencers on virtual social networks extracted from implicit feedbacks are treated as auxiliary resources for recommendation. Empirical results demonstrate the efficacy and effectiveness of our method, particularly, on handling the so-called *cold-start* problem.

Categories and Subject Descriptors

H.2.8 [**Information Systems**]: Database applications—*Data mining*; H.3.3 [**Information Search and Retrieval**]: Information filtering

General Terms

Algorithms, Design, Experimentation

Keywords

news recommendation, matrix factorization, social network, expert

1. INTRODUCTION

Recommending news articles in personalized web services has become a promising research direction with the development of Internet technologies for fast accessing real-time information around the world. Popular news portals, such as Google News and Yahoo! News, have gained increasing attention from a gigantic amount of online news readers. With the large volume of news events happening everyday, an important issue of online news reading services is how to help readers find interesting articles that maximally match their reading appetites, which is called *personalized news recommendation*.

To address the aforementioned problem, most researchers try to develop news recommender systems by utilizing content-based methods (i.e., analyzing the content of news articles to model users' reading preferences) or collaborative filtering (i.e., exploring similar users' reading behaviors) or hybrid version of these two techniques (i.e., combining these two types of methods to alleviate their corresponding drawbacks). Despite a few recent advances, a couple of critical issues of news personalization remain unsolved in previous studies:

- How to overcome the data sparsity problem? Many online users read limited news articles compared with the entire article repository, and therefore the access matrix is very sparse, and the similarity of users' access patterns cannot be effectively captured.
- How to deal with the cold-start problem, including the user cold-start and the item cold-start problems? The former is resulted from the fact that online user groups are evolving, whereas the latter is due to the dynamic nature of news articles.

Many prior research efforts [12, 20] indicate that when there are enough rating records, collaborative filtering approaches can generate better recommendations. However, collaborative filtering approaches fail to extract similar user groups on insufficient data; in such case, researchers turn to explore the potential of "word of mouth" by employing real social networks, e.g. social trust [15, 9]. Since a real social network is usually not readily available, a few researchers model information flow patterns in the virtual social network where people unintentionally influence each other [24]. Prior collaborative filtering and social networking approaches are generally not applicable to new users and new items, since

CIKM'12, October 29–November 2, 2012, Maui, HI, USA.

they have no historical ratings. For new items, the content of the items is usually helpful [6, 3, 4]. For new users, many researchers [16, 22, 26] collect simple user profiles by requiring new users to fill out a list of questionnaires. The questionnaire-based approaches require explicit user inputs. In addition, the quality of user answers can not be guaranteed.

In our work, to resolve the data sparsity and cold-start issues mentioned above, we propose **PRemiSE**, a novel *P*ersonalized news *Re*commendation framework via i*m*pl*i*cit *S*ocial *E*xperts, by incorporating content information, collaborative filtering and information diffusion in virtual social network into probabilistic matrix factorization. Ratings are generated as the aggregation of user preference, semantic item profiles and preferences of the most influential experts. In this way, the semantics of item profiles and the structure of implicit user network are taken into account, leading to better recommendation results. New items are recommended to users who are specifically interested in the associated topics. For new users without any (or enough) ratings and collaborations, items are chosen based on expert opinions in the reading community. In summary, the contribution of our proposed recommendation framework is capable of handling the cold-start problem by leveraging the *opinions* of implicit "experts", and is able to generate better predictions by bridging "word of mouth" and "users with similar tastes".

2. RELATED WORK

Recently, recommending news articles or other document-format web objects has attracted more and more research attentions. Existing news recommender systems can be roughly categorized into three different groups: content-based, collaborative filtering and hybrid methods.

Content-based Recommenders: Content-based recommenders model news readers' reading preferences based on the text content of news items via bag-of-word model or topic models. Such representations are then used as the base to calculate the affinity score between newly-published news articles and the user preference profiles. Representative examples include [1, 6]. Newsjunkie [6], which filters news stories by formal measures of information novelty, shows how the techniques can be used to custom-tailor newsfeeds based on a user's reading history. YourNews [1] intends to increase the transparency of adapted news delivery by allowing users to adapt user profiles. Content-based methods are able to generate reasonable recommendations for new users in content-wise; However in some cases, a bag of words might not be sufficient to capture the reading behaviors of users.

Collaborative Filtering: Collaborative filtering based systems assume that some intrinsic correlations exist among the reading behaviors of users, i.e., if two users are interested in the same topic, they would read similar news articles relevant to this topic. Based on this assumption, most collaborative filtering methods analyze the click behaviors of news readers instead of the news content, either using a group of users "similar" to the given user to predict news ratings [19, 21], or modeling users' behaviors in a probabilistic way [7, 17]. Recently, matrix factorization [12] (MF) and probabilistic matrix factorization [20] (PMF) have revealed to be superior than traditional k-nearest neighbor collaborative filtering methods in other scenarios such as movie recommendation. Extensions of MF & PMF have been published on implicit feedbacks [11] and k-nearest neighbors [10] etc. Collaborative filtering makes a strong assumption that users' behaviors overlap with each other; however, such an assumption might not hold, especially when the item universe changes frequently. In addition, collaborative filtering cannot effectively handle the *cold-start* problem, i.e., a user does not have enough consumption history.

Hybrid Recommenders: In practice, the reading behavior of a news reader can be related to not only the news content, but the reading behaviors of other news readers. Hybrid recommenders aim at combining both content filtering and collaborative filtering to provide more meaningful recommendation. Examples of hybrid recommenders include [3, 4], in which the inability of collaborative filtering to recommend news items is commonly alleviated by combining it with content-based filtering. Recently, researchers start using social network to model online readers' interactions. A couple of algorithms relevant to this [5, 13, 18, 23, 25] have been proposed in the last decade. Researchers also focus on utilizing the trust relations among social media users to derive the recommendation solution[15]. [8] performs a random walk model combining the trust-based and the collaborative filtering approach for recommendation.

Summary: Our proposed model in this paper differs from existing works in the following aspects: (1) In previous works, real social networks (e.g., Epinion's trust network) are required. In this paper, hidden diffusion patterns are modeled on a virtual social network. (2) To the best of our knowledge, only ExpertsCF in [2] applies a set of experts for recommendation. While ExpertsCF is a variation of collaborative filtering using "true experts" obtained from external data source, we automatically detect experts with powerful influence in the semantic community inside the data. (3)We embed content-based, collaborative filtering and social networking approaches into a unified probabilistic framework. We believe such a balanced prediction is consistent to the real decision making process in news consumption.

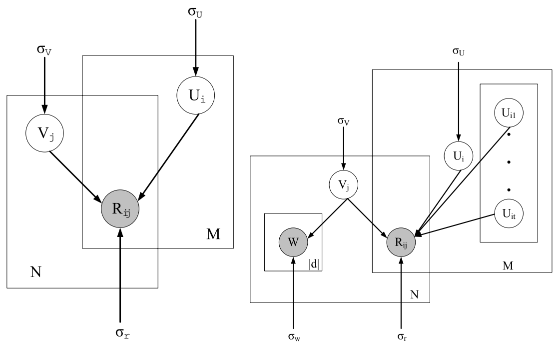
3. PREMISE

3.1 Intuition

In basic matrix factorization models, user-item rating matrix is decomposed to user and item factors. Suppose users' preferences can be represented as $K-$dimensional vectors $\{U_i\}, i \in 1 \sim M$, each of which indicates how much the given user prefers the K underlying factors. Similarly, item coefficients in corresponding factors are denoted by $\{V_j\}, j \in 1 \sim N$. Basic matrix factorization models assume user-item rating is generated from the correlation of factors, i.e., $\hat{r}_{i,j} \sim U_i V_j$. Since most recommendation systems preprocess ratings to be in the range of $[0, 1]$, usually a monotonic function $g(x) = 1/(1 + \exp(-x))$ is adopted to project $U_i V_j$ into the desired scale. A probabilistic expression of this intuition can be defined in Eq.(1), as illustrated in Figure 1(a):

$$p(\hat{r}_{i,j}|U_i, V_j) = \mathcal{N}(\hat{r}_{i,j}|g(U_i V_j), \sigma_r^2). \tag{1}$$

Basic probabilistic matrix factorization models fall into the category of collaborative filtering, and thus it can automatically detect users with similar tastes and items with similar ratings; however, it cannot capture the diffusion of information from person to person. The so-called "word of mouth" is highly valuable in marketing, because individuals



(a) Basic PMF (b) PRemiSE

Figure 1: Graphical Presentation of Basic PMF and PRemiSE.

tend to believe authorities when it comes to choice making. Suppose for a given user U_i, we know a list of "experts" who have significant influence on him/her, then we can borrow the "experts" opinion to predict the given user's reading behavior. We assume influence relationship between "experts" U_e and user U_i is denoted by $\rho_{e,i}$.

Formally, the rating assigned from user U_i to item V_j is then drawn from several experts' opinions, i.e., $\hat{r}_{i,j} \sim \sum_e \hat{r}_{e,j}\rho_{e,i}$. Combining the two probabilistic generation processes, we have:

$$p(\hat{r}_{i,j}) = \mathcal{N}(\hat{r}_{i,j}|g(V_j(w_1U_i + w_2\sum_e U_e\rho_{e,i})), \sigma_r^2), \qquad (2)$$

where the parameters w_1 and w_2 coordinate the effect of collaborative filtering and "word of mouth", subject to $w_1 + w_2 = 1$. Note that w_1 and w_2 are dynamic parameters. Intuitively, for experienced users, their preferences dominate their decision; while for unexperienced users, they are more likely to follow the experts.

The content of news also contains rich information. In general, word distribution is factor-dependent. Suppose Θ is the word space, where $\theta_{w,k}$ is the strength of word w correlating with factor k. We associate the TFIDF weighting of terms in each item profile with item factor and Θ. Probably each item j's profile $\vec{\mathbf{w}}_j$ is a Gaussian distribution whose mean is the product of item factors V and Θ, i.e.,

$$p(\hat{\vec{\mathbf{w}}}_j|V_j, \Theta) = \mathcal{N}(\hat{\vec{\mathbf{w}}}|g(\Theta V_j), \sigma_w^2), \qquad (3)$$

where U, V, Θ are zero-mean spherical Gaussian priors, i.e.,

$$p(\Theta) = \Pi_{w,z}\mathcal{N}(\Theta_{w,z}|0, \sigma_w^2 I),$$
$$p(U) = \Pi_{i,z}\mathcal{N}(U_{i,z}|0, \sigma_u^2 I),$$
$$p(V) = \Pi_{j,z}\mathcal{N}(V_{j,z}|0, \sigma_v^2 I).$$

By combining collaborative filtering, "word of mouth" in social networks and content semantics, PRemiSE is illustrated in Figure 1(b).

3.2 Inference

Suppose we have learned U^*, V^*, Θ^* from the training set, then in PRemiSE we perform inference to predict ratings $r_{i,j}$ in the test set. We consider the following four cases:

1. **Existing Item by Existing User:** If the user U_i is active and has sufficient ratings in the training set, the item V_j is frequently accessed in the training set, then we will have accurate estimation of user factor U_i and item factor V_j. In this case, rating prediction will be the expectation of the generated rating from collaborative filtering and expert opinions.

$$r_{i,j} = g(V_j(w_1U_i + w_2\sum_e U_e\rho_{e,i})). \qquad (4)$$

2. **Existing Item by New User:** If the target user i is new or "lazy" in the training set, i.e., the user has limited reading history, then we may not have available user factor U_i. We consult to the experts of the community. We denote influence power of expert U_e to a random "null" user ϕ as $\rho_{e,\phi}$, then we use the following Eq.(5) to predict rating. Note that Eq.(5) can also be treated as a special case of Eq.(4) when $w_1 = 0, w_2 = 1$.

$$r_{i,j} = g(V_j(\sum_e U_e\rho_{e,\phi})). \qquad (5)$$

3. **New Item by Existing User:** For any incoming news item or an unpopular news V_j without enough ratings, we can approximate its factors based on its content $\vec{w}_j$. Note that in Eq.(4), the item vector V_j can be factorized as $\vec{w}_j\Theta^{-1}$. Then we can transfer this problem as

$$r_{i,j} = g(\vec{w}_j\Theta^{-1}(w_1U_i + w_2\sum_e U_e\rho_{e,i})). \qquad (6)$$

4. **New Item by New User:** To predict how a new user U_i will rate a new item V_j, we first compute pseudo ratings from global experts to V_j, then we aggregate pseudo ratings from all experts as in Eq.(7). Note that this is a special case of Eq.(4), where the profile of the new item V_j is factorized as $\vec{w}_j\Theta^{-1}$, and the fusion coefficient of user interest is decreased by setting $w_1 = 0, w_2 = 1$.

$$r_{i,j} = g(\vec{w}_j\Theta^{-1}(\sum_e U_e\rho_{e,\phi})). \qquad (7)$$

4. EXPERIMENTAL EVALUATION

4.1 Real-World Dataset

The news data set is crawled from several popular news service websites, ranging from July 15th, 2010 to Dec 16th, 2010 [14]. It contains the details of news articles (e.g., news title, content, published time, etc.) and user access history (e.g., anonymous users, accessed news items, accessed time, etc.). In order to investigate the feasibility of our "expert" model, we consider two types of elements for constructing implicit social graph: news stories and named entities. In the experiment, news stories are formed by concatenating news title and news content, whereas named entities are extracted from news stories via information extraction tools as in [14].

For evaluation purpose, we divide the entire news data into two disjoint sets, where the first one ranges from July 15th, 2010 to September 30th, 2010, and the remaining falls into the second set. We then construct implicit social networks for these two sets using the two types of elements, respectively. Hence we can obtain 4 different data sets. Note

Table 1: Statistics of news data

Type	Story				Entity			
DataSet	DataSet1		DataSet2		DataSet3		DataSet4	
Split	Training	Test	Training	Test	Training	Test	Training	Test
#Users	1,583	550	12,022	540	41,699	8,606	292,274	22,119
#Items	772	423	6,426	514	8,195	1,998	44,114	3,584
#terms	138,999		236,174		187,281		228,462	
#Ratings	29,255	9,367	347,919	9,563	513,488	152,001	5,246,968	179,844
#new User	-	138	-	68	-	6,532	-	15,674
#new Item	-	202	-	208	-	360	-	278
TimeSpan	07|15-09|28	09|29-09|31	10|01-12|13	12|14-12|16	07|15-09|28	09|29-09|31	10|01-12|13	12|14-12|16

that in each data set, news stories associated with users' access histories of the last 3 days are extracted to serve as the test data, whereas the remaining is treated as the training data. In dataset4, each user has rated at least 30 items (entities) in the training set, and each item has been accessed for more than 30 times. In the other three data sets, each user has at least 10 items in the training set, and each item has been rated by at least 10 individuals. The ratings in the Story data sets are binary. The ratings in the Entity data sets are numerical, which are the normalized accumulative visiting number by each user on each entity. The statistics of the four data sets is depicted in Table 1.

4.2 The Comparative Study

Benchmarks used in this comparative study include three recommendation algorithms, namely collaborative filtering (CF) [19], basic matrix factorization (MF) [12], and Latent Dirichlet Allocation filtering (LDA) [22]. In CF, the predicted rating $\hat{r}_{i,j}$ from user i to item j is derived from i's similar users' ratings on j, i.e.,

$$\hat{r}_{i,j} = \frac{\sum_k sim(i,k) r_{k,j}}{\sum_k sim(i,k)}, \quad (8)$$

where the similarity between user k and i is defined as the cosine similarity over ratings. In basic matrix factorization, the predicted rating $\hat{r}_{i,j}$ is generated by Eq.(1).

LDA is originally proposed to model topics in documents. In the experiments, we employ LDA to estimate the distribution over K topics of each item profile. Then we predict the probability of the given user to generate a given item, i.e.,

$$p(V_j|U_i) = \sum_z^K p(V_j|z)p(z|U_i). \quad (9)$$

Here $p(z|V_j), p(z|U_i)$ can be obtained by LDA. Using Bayersian rule, we have

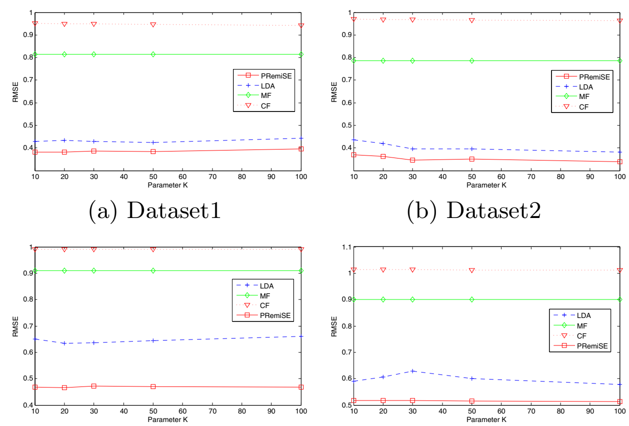
$$p(V_j|z) = \frac{p(z|V_j)p(V_j)}{p(z)} = \frac{p(z|V_j)p(V_j)}{\sum_V' p(z|V')p(V')}, \quad (10)$$

where we can assume $p(V') = 1/N$. Intuitively, if the conditional probability by Eq.(9) is significantly larger than random, then the predicted rating should be definite, i.e.,

$$\hat{r}_{i,j} = \begin{cases} 1 & \text{if } p(V_j|U_i) > 2/N, \\ \frac{N \times p(V_j|U_i)}{2} & \text{otherwise.} \end{cases}$$

There is no mechanism in MF, CF and LDA-filtering to deal with the *cold-start* problem. We randomize a user factor in MF and LDA-filtering as a pseudo user, in CF the average rating is assigned as the predicted rating for new users and new items. In PRemiSE, the dynamic parameters w_1, w_2 are set to be $w_1 = 1, w_2 = 0$ for users and items that appear in the training set, because after pre-processing users and items in the training set have enough ratings. We assume that they will not be affected much by local experts. The evaluation metric is RMSE, which measures the root mean square error from the predicted rating and the actual rating in the test sets. It can be calculated as

$$RMSE(\hat{R}, R) = \sqrt{\frac{\sum(\hat{r}_{i,j} - r_{i,j})^2}{|R|}}. \quad (11)$$



(a) Dataset1 (b) Dataset2

(c) Dataset3 (d) Dataset4

Figure 2: Comparative Study over parameter K in 4 data sets.

As shown in Figure 2, the RMSEs of the 4 methods on parameter K are presented for each data set. On all the data sets, collaborative filtering has maintained the worst performance, due to the sparsity of the data sets. The access frequency for news is relatively small compared with other types of items, e.g. music, movie, etc. Therefore standard collaborative filtering is not able to produce similar users. When we use entities as items, the performance of collaborative filtering is slightly worse than using news articles, resulting from the less density of the entity data.

The performance of MF is better than collaborative filtering, but is slightly worse than LDA-filtering. Note that LDA-filtering is capable of predicting user's preference on new items. Our analysis shows that when predicting existing items for existing users, LDA-filtering is not better than MF. But it is superior in predicting new items for existing users. This suggests that in scenarios like news recommendation, the rich information contained in item profiles can be very helpful.

PRemiSE performs best on all data sets. The best K is different among each data set. For example, in dataset1, the best performance is achieved when $K = 20$, and in dataset2 it is $K = 100$. This indicates that for more complex data sets, more factors are required to capture the semantic difference between items.

4.3 Handling the Cold-start Problem

We now look further on how different strategies work for the cold-start problem. We compute RMSE on dataset1 for four cases: predict ratings for existing users (users exist in training set) to existing items (items exist in training set), predict ratings by existing users to new items (items that do not exist in training set), predict ratings by new users to existing items, and predict ratings by new users to new items. We abbreviate these four cases as oop, onp, nop, and nnp.

As mentioned above, MF, LDA-filtering and CF are not capable of handling all these cases. Our modifications are as follows: for onp, MF randomizes an item factor, and generates ratings from the user factor and the pseudo item factor; LDA-filtering inferences the topic distribution for the new item, and generates ratings as usual. For nop, MF randomizes a user factor, and generates ratings from the pseudo user factor and the item factor; LDA-filtering randomizes the topic distribution for the new user, and generates ratings by the pseudo user topics and the item topics. For nnp, MF and LDA-filtering give a random guess. CF uses the average rating for onp, nop and nnp.

The RMSE performance of the four algorithms on the cold start problem and all rating predictions are shown in Figure 3. We can observe that PRemiSE is comparable with MF in existing item and existing user rating prediction, but outperforms MF and LDA-filtering in predicting cold start problems, including new users and new items.

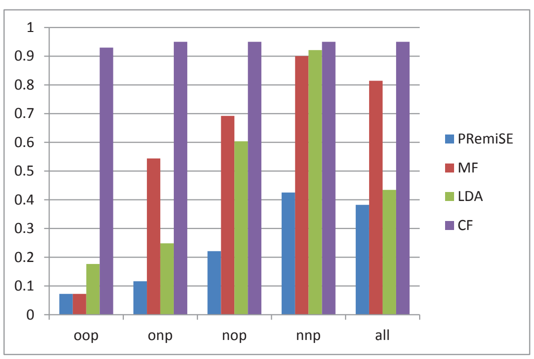


Figure 3: Comparative study on cold start problem.

Acknowledgement

Chen Lin is partially supported by Shanghai Key Laboratory of Intelligent Information Processing under grant No. IIPL-2011-004, China Natural Science Foundation under grant No.61102136,61001013, the Natural Science Foundation of Fujian Province of China under grant No.2011J05158), the Fundamental Research Funds for the Central Universities under Grant No.2011121049. Lei Li and Tao Li are partially supported by US NSF grants DMS-0915110 and CNS-1126619 and by the Army Research Office under grant number W911NF-10-1-0366.

5. REFERENCES

[1] J. Ahn, P. Brusilovsky, J. Grady, D. He, and S. Syn. Open user profiles for adaptive news systems: help or harm? In *WWW*, pages 11–20, 2007.

[2] X. Amatriain, N. Lathia, J. M. Pujol, H. Kwak, and N. Oliver. The wisdom of the few: a collaborative filtering approach based on expert opinions from the web. In *SIGIR*, pages 532–539, 2009.

[3] R. Burke. Hybrid systems for personalized recommendations. *Intelligent Techniques for Web Personalization*, pages 133–152, 2005.

[4] W. Chu and S. Park. Personalized recommendation on dynamic content using predictive bilinear models. In *WWW*, pages 691–700, 2009.

[5] S. Debnath, N. Ganguly, and P. Mitra. Feature weighting in content based recommendation system using social network analysis. In *WWW*, pages 1041–1042, 2008.

[6] E. Gabrilovich, S. Dumais, and E. Horvitz. Newsjunkie: providing personalized newsfeeds via analysis of information novelty. In *WWW*, pages 482–490, 2004.

[7] T. Hofmann. Latent semantic models for collaborative filtering. *ACM Transactions on Information Systems*, 22(1):89–115, 2004.

[8] M. Jamali and M. Ester. Trustwalker: a random walk model for combining trust-based and item-based recommendation. In *SIGKDD*, pages 397–406, 2009.

[9] M. Jamali and M. Ester. Using a trust network to improve top-n recommendation. In *RecSys*, pages 181–188, 2009.

[10] Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *SIGKDD*, pages 426–434, 2008.

[11] Y. Koren. Factor in the neighbors: Scalable and accurate collaborative filtering. *ACM Transactions on Knowledge Discovery from Data.*, 4:1:1–1:24, January 2010.

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

[13] K. Lerman and R. Ghosh. Information contagion: An empirical study of the spread of news on digg and twitter social networks. In *ICWSM*, 2010.

[14] L. Li, D. Wang, T. Li, D. Knox, and B. Padmanabhan. Scene: a scalable two-stage personalized news recommendation system. In *SIGIR*, pages 125–134, 2011.

[15] H. Ma, I. King, and M. Lyu. Learning to recommend with social trust ensemble. In *SIGIR*, pages 203–210, 2009.

[16] S.-T. Park and W. Chu. Pairwise preference regression for cold-start recommendation. In *RecSys*, pages 21–28, 2009.

[17] D. Pennock, E. Horvitz, S. Lawrence, and C. Giles. Collaborative filtering by personality diagnosis: A hybrid memory- and model-based approach. In *UAI*, pages 473–480, 2000.

[18] O. Phelan, K. McCarthy, and B. Smyth. Using twitter to recommend real-time topical news. In *RecSys*, pages 385–388, 2009.

[19] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. GroupLens: an open architecture for collaborative filtering of netnews. In *CSCW*, pages 175–186, 1994.

[20] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. *NIPS* , 20:1257–1264, 2008.

[21] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl. Item-based collaborative filtering recommendation algorithms. In *WWW*, pages 285–29, 2001.

[22] Y. Seroussi, F. Bohnert, and I. Zukerman. Personalised rating prediction for new users using latent factor models. In *HT'11*, pages 47–56, 2011.

[23] A. Seth and J. Zhang. A social network based approach to personalized recommendation of participatory media content. In *ICWSM*, 2008.

[24] X. Song, B. L. Tseng, C.-Y. Lin, and M.-T. Sun. Personalized recommendation driven by information flow. In *SIGIR*, pages 509–516, 2006.

[25] J. Wang, Q. Li, Y. Chen, J. Liu, C. Zhang, and Z. Lin. News recommendation in forum-based social media. In *AAAI*, pages 1449–1454, 2010.

[26] K. Zhou, S.-H. Yang, and H. Zha. Functional matrix factorizations for cold-start recommendation. In *SIGIR*, pages 315–324, 2011.