Multi-source Information Fusion for Personalized Restaurant Recommendation

Jing Sun
Shanghai Key Laboratory of Data Science
School of Computer Science
Fudan university
jingsun@fudan.edu.cn

Yun Xiong
Shanghai Key Laboratory of Data Science
School of Computer Science
Fudan university
yunx@fudan.edu.cn

Yangyong Zhu
Shanghai Key Laboratory of Data Science
School of Computer Science
Fudan university
yyzhu@fudan.edu.cn

Junming Liu
Rutgers Business School
Rutgers University
liujm000@gmail.com

Chu Guan
University of Science and Technology of China
guanchu0329@gmail.com

Hui Xiong
Rutgers Business School
Rutgers University
xionghui@gmail.com

ABSTRACT

In this paper, we study the problem of personalized restaurant recommendations. Specifically, we develop a probabilistic factor analysis framework, named RMSQ-MF, which has the ability in exploiting multi-source information, such as the users' task, their friends' preferences, and human mobility patterns, for personalized restaurant recommendations. The rationale of this work is motivated by two observations. First, people's preferences can be affected by their friends. Second, human mobility patterns can reflect the popularity of restaurants to a certain degree. Finally, empirical studies on real-world data demonstrate that the proposed method outperforms benchmark methods with a significant margin.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information filtering

Keywords

Restaurant Recommendation, Matrix Factorization, Bayesian models, Mobile Computing

1. INTRODUCTION

Recent years have witnessed a rapid development on restaurant recommendation and advertisements. With the emergence of social media and mobility applications, a more comprehensive and effective restaurant recommendation strategy provides the opportunity to offer convenience to customers as well as to improve profits for restaurants. The largest social media of restaurant, Yelp,[1] have provided a set of benchmark 5-point rating for restaurants in US, expressing the extent of the customers' favors on each restaurant. Previous work [3] exploited the multi-aspect ratings of restaurant to uncover the user preference. However, without considering the social factor and mobility factor, the recommendation accuracy is not high enough and need further improvement.

Indeed, traditional recommender systems usually assume that the user and the item data are independent and identically distributed, which ignores the interrelationships among items. As users' preferences are not actually binary decisions and have a certain granularity, manually specifying personal preferences is obstructive and usually bring a biased recommendation to users. To the best of our knowledge, limited efforts have been made for understanding customers' preferences for restaurants. Indeed, the tradition personalized ranking targets at building ranking models to disclose the user interests from profiles of rating event, such as user-dependent features, item-dependent features, and shared features. Also, filtering based methods can be divided into two categories: collaborative filtering and content-based filtering. Collaborative filtering recommends items to the users whose tastes are similar whereas content-based filtering recommends items similar to those the user has liked in the past. Recently, matrix factorization models in collaborative filtering become more and more popular. Matrix factorization assumes users' preferences can be generated by user latent factors and item latent factor in a latent space.

However, restaurants interact with people through daily dining. There are several unique characteristics of restaurant review systems, which make traditional matrix factorization techniques difficult to be adapted for restaurant recommendation. In fact, the decision process of a user choosing a restaurant is more complicated because of so many influence factors: 1) users tend to check around several centers, where the check-in locations follow Gaussian distribution at each center for a typical user's rating behavior. Second, users can be easily influenced by the friends they trust, and prefer their friends' recommendations. 2) two restaurants with similar or the same semantic topics can have different popularities if they are located in different regions. 3) popularity affects purchase behaviors. The probability of a

[1]http://www.yelp.com

SIGIR '15 August 09 - 13, 2015, Santiago, Chile
Copyright 2015 ACM ISBN 978-1-4503-3621-5/15/08 ...$15.00
DOI: http://dx.doi.org/10.1145/2766462.2767818.

user choosing a restaurant is partly affected by the traffic conditions around the restaurant. 4) the users have dynamic mobility behaviors, which impose the challenges on the modeling of restaurant decision. While there are some studies on restaurant recommendations [3, 2], it is necessary to build an integrated analysis of the joint effect of the factors mentioned above. Intuitively, if a flexible method is utilized to collect more information, then better performance may be achieve. Therefore, we need a model that jointly encodes the user ratings, social factors, region topic and mobility factors into the user decision process to learn user preferences for effective recommendation recommendations.

In this paper, a Bayesian graphical model taking account of both collaborative filtering latent factors, the social factor and mobility factor is proposed to learn the user preferences. We carried out the experiments to show the efficiency and effectiveness of the proposed approach with real world data evaluations.

2. RELATED WORK

Content-based approaches and collaborative filtering (CF) approaches have been the prevailing recommendation strategies. Content-based algorithms are based on a description of the item and profile of the user's preference, which try to recommend items according to a user's historical preference. CF algorithms assume that people's behavior is affected by someone with similar preferences, and user's ratings are adjusted by his neighbors' ratings or recommendation. CF approaches can be categorized into two classes: memory-based CF and model-based CF. Memory-based CF algorithms first search for neighbors who have similar rating histories to the target user. Then the target user's rating histories can be predicted according to his neighbors' ratings. Model-based CF algorithms use various models and machine learning techniques such as clustering algorithms or Bayesian networks to discover latent factors that account for the observed ratings. A particular type of CF algorithm uses matrix factorization, a low-rank matrix approximation technique [4].

In contrast, restaurant recommender systems can also be developed in terms of users' past preferences via matrix factorization (MF), such as SVD++ [8], PMF [6], NMF [1]. One drawback of MF is the so-called ramp-up problem, in which MF performs poorly when the recommender system is not initialized with a sufficient number of ratings. Therefore, more studies such as BiasedMF [8] and SocialMF [5] incorporated side and prior information into latent factor models by exploring users' and items' observable features. The work [7] studied on restaurant recommendation relied on knowledge about users and restaurants to generate recommendations. However, previous studies in restaurant recommendation still have several limitations: they lack of integrated modeling of the joint influence of personalization in user ratings, social factors, geographic interrelationship between users and restaurants, bias effect and profile similarity.

3. PROPOSED MODEL

3.1 Problem Definition

The problem of restaurant recommendation is to recommend restaurant to a user based on ratings and other side information (e.g.,user profile, social factor and mobility factor information) collected from a web site. Let $U = \{u_1, u_2, \cdots, u_u, \cdots, u_p\}$ be a set of users, where each user has a user profile (e.g.location, yelping time). Let $V = \{v_1, v_2, \cdots, v_i, \cdots, v_q\}$ be a set of restaurants, where each restaurant has a location $l_i = \{lon_i, lat_i\}$ in terms of longitude and latitude, and observable restaurant-dependent features (e.g.business hours, cuisine categories, etc.) which describe the restaurant features. We denote the rating as R_{ui} for the pair of user u and restaurant i. We also have social influence weight between two friends that is based on both of their social connections and similarity of their visiting activities. $S_{uf} = \delta\frac{|\tau(u)\cap\tau(f)|}{|\tau(u)\cup\tau(f)|} + (1-\delta)\frac{|A_u\cap A_f|}{|A_u\cup A_f|}$, where δ is a tuning parameter ranging within [0,1], and $\tau(u)$ and A_u denote the friend set and visiting restaurants set of user u, respectively. Also, we refer u as user and i as restaurant in the following unless otherwise specified.

3.2 Model specification

Figure 1 shows the decision process of user u to choose restaurant i in a generative way. RSMQ-MF model fuses personal interest in restaurants(Q), personal interest in the area around restaurants(M) and social influence(S). The main contribution of the approach is that we fuse the information of the region around the restaurants and use the taxi drop-off as the popularity of the region. Here, we firstly infer the proposed personalized recommendation model to predict users' ratings. And then, we show the training method of our approach. For the user and restaurant latent feature on

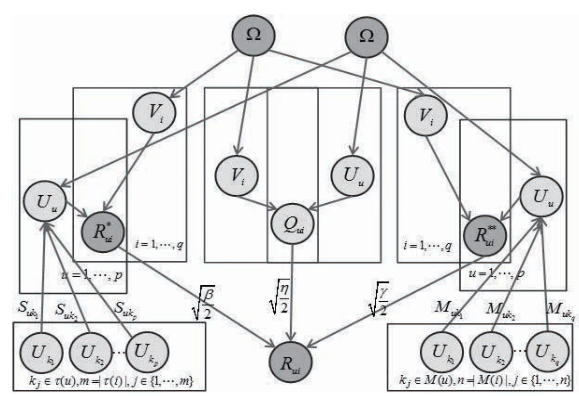


Figure 1: The graphical representation of the proposed RSMQ-MF model.

the personal interest, we have the conditional distribution according to:

$$p(Q|U, V, \Omega) = \prod_u \prod_i [\mathcal{N}(Q_{u,i}|U_u, V_i^\top, \Omega)]^{I^R_{u,i}} \tag{1}$$

where $\mathcal{N}(|\mu, \Omega)$ is the gaussian distribution with mean μ and variance Ω. $I^R_{u,i}$ is the indicator function which is equal to 1 if user u has rated restaurant i and equal to 0 otherwise. Note that the condition distribution based on the users' ratings is similar with the conditional probability based on personal latent interest:

$$p(R|U, V, \Omega) = \prod_u \prod_i [\mathcal{N}(R_{u,i}|U_u, V_i^\top, \Omega)]^{I^R_{u,i}} \tag{2}$$

Given the latent features on the topic interest of his friends, for the user latent feature based on interest topic, we have the conditional distributions as follows:

$$p(U|M,\Omega)=\prod_{u}\mathcal{N}(U|\sum_{f\in U}M_{u,f}U_f,\Omega) \tag{3}$$

Here, f is a friend of user u. As mentioned above, there are three main aspects influencing personalized recommendation in our model: 1) social influence $S_{u,f}$, which means you would trust your friends; 2) interest topic $M_{u,f}$, which means your region interest topic is similar with your friends', which is probability affected by the taxi drop-off information (see in the Experiment section); 3) personal interest $Q_{u,i}$, which means the effect on the restaurants you are interested based on user and restaurant profiles. Thus, by combining the social influence S, interest similarity M of the friends and user personal interest Q with the observed ratings R , the proposed RSMQ-MF model decreases the training error. According to the basic MF model, through bayesian inference, the posterior probability of our model is as follows:

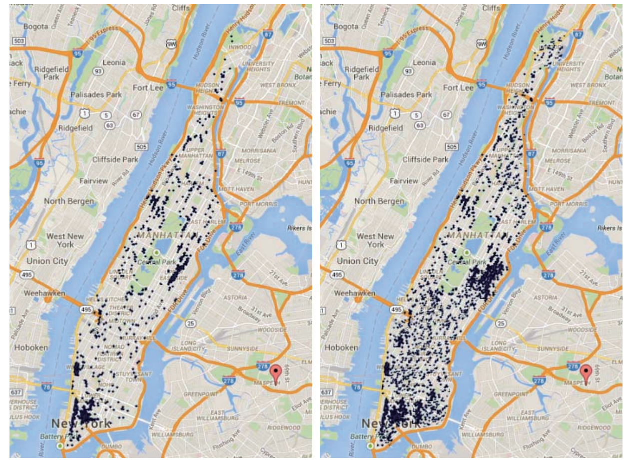
$$\begin{aligned}
&p(U,V|R,S,M,Q,\omega)\propto p(R|U,V,\Omega)p(U|S,\Omega)p(U|M,\Omega)\\
&p(Q|U,V,\Omega)\\
&p(U|\Omega)p(V|\Omega)=\prod_{u}\prod_{i}[\mathcal{N}(R_{u,i}|U_u,V_i^{\top},\Omega)]^{I^R_{u,i}}\times\\
&\prod_{u}\mathcal{N}(U|\sum_{f\in U}S_{u,f}U_f,\Omega)\times\prod_{u}\mathcal{N}(U|\sum_{f\in U}M_{u,f}U_f,\Omega)\\
&\prod_{u}\prod_{i}[\mathcal{N}(Q_{u,i}|U_u,V_i^{\top},\Omega)]^{I^R_{u,i}}\times\\
&\prod_{u}\mathcal{N}(U_u|0,\Omega)\times\prod_{i}\mathcal{N}(V_i|0,\Omega)
\end{aligned} \tag{4}$$

Keeping the prior variance and observation noise fixed, we have the loss function:

$$\begin{aligned}
&\mathcal{L}(R,U,V,S,M,Q)=\frac{1}{2}\sum_{u,i}(R_{u,i}-\hat{R}_{u,i})^2+\frac{\lambda}{2}(\|U\|_F^2+\\
&\|V\|_F^2)+\frac{\beta}{2}\Sigma_u((U_u-\Sigma_f S_{u,f}U_f)(U_u-\Sigma_f S_{u,f}U_f)^{\top})\\
&+\frac{\gamma}{2}\Sigma_u((U_u-\Sigma_f M_{u,f}U_f)(U_u-\Sigma_f M_{u,f}U_f)^{\top})\\
&+\frac{\eta}{2}\Sigma_{u,i}(Q_{u,i}-U_uV_i^{\top})^2
\end{aligned} \tag{5}$$

Here, $\hat{R}_{u,i}$ is the predicted rating. $\|.\|_F^2$ denotes Frobenius norm. λ is regularization weight, and β, γ and η are iteration step size. In Eq.(5), the idea of social influence is the second term, which means that user personal interest would be similar with the friends latent interest. The third term is enforced as the similarity of the interest topic. And the factor of user personal interest is the last term which means that users would be interested with several restaurant. To obtain a local minimum of the objective function given by Eq.(5), we perform gradient descent in U_u and V_i:

$$\begin{aligned}
&\frac{\partial\mathcal{L}}{\partial U_u}=\sum_{i}I^R_{u,i}(\hat{R}_{u,i}-R_{u,i})V_i+\lambda U_u+\beta(U_u-\\
&\Sigma_v S_{u,f}U_f)-\beta\Sigma_{f:u\in F_f}S_{u,f}(U_f-\Sigma_{\omega\in F_f}S_{v,w}U_w)+\\
&\gamma(U_u-\Sigma_{f\in F_u}M_{u,f}U_f)-\gamma\Sigma_{f:u\in F_f}M_{f,u}(U_f-\\
&\Sigma_{\omega\in F_f}M_{f,\omega}U_\omega)+\eta\Sigma_i I^R_{u,i}(U_uV_i^{\top}-Q_{u,i})V_i
\end{aligned} \tag{6}$$



(a) restaurants (b) taxis

Figure 2: A demonstration of restaurants and taxis drop-off points in the dataset.

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial V_i}&=\sum_{u}I^R_{u,i}(\hat{R}_{u,i}-R_{u,i})U_u+\lambda V_i\\
&+\eta\sum_{u}I^R_{u,i}(U_uV_i^{\top}-Q_{u,i})U_u
\end{aligned} \tag{7}$$

where $I^R_{u,i}$ is the indicator function that is equal to 1 if user u has rated restaurant i, and 0 otherwise. The initial values of U and V are sampled from the Gaussian distribution with zero mean. The latent feature matrix learning was empirically affected. We adjust the step size based on the previous values of the two latent feature matrix U and V in each iteration.

4. EXPERIMENTS

4.1 Date set and Evaluation Metrics

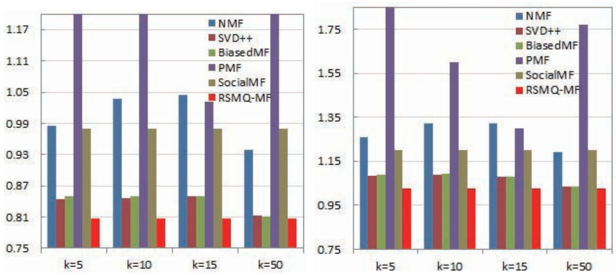
Data set. We crawled the data from yelp.com site covering Manhattan district. We collect a complete snapshot, including users' profile, user's rating, users' friend list, and restaurants' profile, from Yelp. The total number of restaurant is 7,115. In order to easy to do experiments to validate the performance of our model. We positively pick the restaurant to limit other type of potential impact. And we select 1,000 restaurants using a greedy algorithm with max distance, which keep we don't have nearest restaurants. The detailed statistics of the final dataset are shown in Table 1. Density is the percentage of entries in the user-restaurant matrix that have ratings.

Moreover, to measure the popularity in Manhattan, we adopt the taxi drop-off information from New York City taxi Data[2] which was collected from Jan. 2013 to Jun. 2013. Setting lunch time period as 11:00-13:30, and dinner time period as 17:30-20:00 in Manhattan. Then, match the restaurant location and taxi customer drop-off longitude and latitude distance, select those restaurant recently and the distance between two is less than 10 meters. Get 21,684,273 taxi drop-off records related to the 1000 restaurants. Figure 2

[2]https://uofi.box.com/NYCtaxidata

Table 1: Size of dataset

# Users	# Restaurants	# Ratings	% Density
1500	1000	33720	2.248%



(a) MAE (b) RMSE

Figure 3: The average MAE and RMSE values at different latent dimensions k.

is a map illustration of restaurants and taxis in Manhattan Area. Each black point representing one restaurant in Figure 2(a) or one taxi drop-off location in Figure 2(b). This map shows there are lots of taxi drop-off behaviors around the restaurant of Manhattan.

Evaluation Metrics. We use the following two metrics for evaluation.

(1) the Mean Absolute Error,
MAE $= \sum_{u,i} |r_{ui} - \hat{r}_{ui}|/N$,

(2) the Root Mean Square Error,
RMSE $= \sqrt{\sum_{u,i} (r_{ui} - \hat{r}_{ui})^2/N}$,

where r_{ui} and $\hat{r}_{ui}$ denote the observed rating and the predicted rating, and N denotes the total number of the tested data. The smaller the value of MAE or RMSE, the more precise a recommendation.

4.2 Performance Comparison

To show the effectiveness of the proposed RSMQ-MF model, we compare the performances of our model against the following baseline algorithms: 1)Nonnegative Matrix Factorization (NMF) [1]. 2)SVD++ [8], which directly incorporates implicit feedback into the singular value decomposition (SVD) model. 3)BiasedMF [8], which is a multifaceted collaborative filtering model. 4)Probabilistic Matrix Factorization (PMF)[6]. 5)SocialMF [5], which is a matrix factorization technique with trust propagation for recommendation in social networks. We randomly selected 80% of data from our dataset as training data and different number of latent factors (k=5,10,15,50) to test all the methods. We set learning rate as 0.01, regularization $\lambda_U = 0.01$ and $\lambda_V = 0.05$ for all methods. For our approach, we set $\beta = 20$, $\gamma = 20$,$\eta = 20$.

From the Figure 3(a), it is clear that our approach outperforms the other methods on different number of latent factors (k). The MAE values of our approach (with average 0.79) are lower than the MAE values of SVD++, BiasedMF, NMF, SocialMF or PMF (with average 1.03). In addition, in all the baselines, BiasedMF and SVD++ works better than NMF, SocialMF and PMF in all k value settings. PMF gives comparatively low performance with 2.219 MAE(k=5), since the model is a pure probabilistic factor model and does not take advantage of social or mobility factors.

Figure 3(b) shows the RMSE comparisons: our approach outperforms all the completing models. For example, the average RMSE of five baselines is about 1.28 whereas the RMSE of our method is only 0.98. SocialMF works better than NMF and PMF because of its aggregation of user social links. The present method not only consider social factors, but also incorporates taxi drop-off information, which makes our approach performing more effective recommendation method.

5. CONCLUSION

In this paper, in order to provide more accurate and efficient restaurant recommendation, we propose a novel fused matrix factorization framework to take into account multi-source information. To incorporate the side information, we explored the bias effect which was regressed from user profiles and restaurant characteristics, as well as exploited the similarity from social factors and mobility factors between users and restaurants. Our model not only discovered the social influence of user by flexibly collecting information from user's social network, but also incorporates the taxi drop-off information into the user preference indices. Experimental results on real-world data shows the proposed RSMQ-MF model can achieve significantly better performance than other state-of-the-art methods.

6. ACKNOWLEDGEMENTS

This work was supported in part by grants from the National Natural Science Foundation of China under Grant No.61170096,71331005.The work was also partially supported by grants from Shanghai Foundation for Development of Science and Technology under Grant No.13dz2260200 13511504300, 14511107302.

7. REFERENCES

[1] D.D.Lee and H. Seung. Algorithms for non-negative matrix factorization. *In Advances in Neural Information Processing Systems*, 20:1257–1264, 2001.

[2] F.Yanjie, L.Bin, G.Yong, and X.Hui. User preference learning with multiple information fusion for restaurant recommendation. In *SDM2014 Conference Proceedings*, pages 470–478, April 2014.

[3] H.H.Lee and G.W.Teng. Incorporating multi-criteria ratings in recommendation systems. In *IRI2007 Conference Proceedings*, pages 273–278, August 2007.

[4] I.Markovsky. *Low-Rank Approximation: Algorithms, Implementation, Applications.* Springer ISBN 978-1-4471-2226-5, 2012.

[5] M.Jamali and M.Ester. A matrix factorization technique with trust propagation for recommendation in social networks. In *RecSys2010 Conference Proceedings*, pages 135–142, Septmber 2010.

[6] R.Salakhutdinov and A.Mnih. Probabilistic matrix factorization. *In Advances in Neural Information Processing Systems*, 13:556–562, 2008.

[7] S.Trewin. Knowledge-based recommender systems. *Encyclopedia of Library and Information Science*, 69(32):180–200, 2000.

[8] Y.Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *KDD2008 Conference Proceedings*, pages 426–434, August 2008.