Hybrid Rating Prediction using Graph Regularization from Network Structure

Group 45

Yilun Wang
Department of Computer Science
Stanford University
yilunw@stanford.edu

Guoxing Li
Department of Computer Science
Stanford University
guoxing@stanford.edu

Jingrui Zhang
Department of Computer Science
Stanford University
jingrui@stanford.edu

***Abstract*—To understand users' preference and business performance in terms of customer interest, rating prediction systems are wildly deployed on many social websites. These systems are usually based on the users past history of reviews and similar users. Such traditional recommendation systems[1] suffer from two problems. The first is the cold start problem. In many cases, the system has no knowledge about the interests of a new user. Second, the data sparsity problem, some users may have a few business ratings. However, a noticeable trend in social life is that people sharing characteristics and attributes tend to stay friends so that knowledge about the person's neighborhood structure demonstrates his preference in certain degree. The traditional recommendation approaches fail to take network structure of users into consideration. To address the aforementioned three problems, we proposed a hybrid method based on network structure and collaborative filtering approach. Specifically, a graph regularization is learned by random walk algorithm which discovers the underlying network structure and the matrix factorization based collaborative filtering is used to predict user ratings. The experiments conducted on Yelp user-business reviews dataset demonstrate the advantages of our proposed method beyond baselines and validate the contribution of its two components.**

Keywords—Social Network, Data Mining, Model, Algorithm, Recommendation System, Hybrid method

I. Introduction

Online reviews have become a popular way for customers to voice their opinion and also one of the most trusted sources of consumer confidence in making decisions. As a result, accurate prediction of star rating for one of the most widely used business ranking sites in the country - Yelp could have huge potential and value. Thus in this paper we approach this task.

Yelp[2] provides an online platform to let Yelpers to evaluate their experience on local businesses. The most important key factors of evaluation are ratings and reviews. Related study has shown that a one-star increase in Yelp rating leads to a 5-9 percent increase in revenue. On the other hand, rating is the most straightforward and direct tool to help Yelpers make a choice. Thus, predicting accurate rating for a business could be an invaluable information for both parties on Yelp. In addition, being able to predict personalized ratings for users would make recommendation on local businesses possible. An accurate recommendation would not only delight user's experience but also drive up traffics of businesses. At last, personalized rating prediction could also benefit search ranking, which adds core value to Yelp's product as search is one of the most commonly used functionalities on the platform.

One of the most popular algorithms used on recommendation is collaborative filtering. There are two problems that traditional collaborative filtering is suffering. One is its performance degrades when data gets sparser (e.g., less historic review data in large number of users/businesses dataset). Matrix factorization resolves this issue really well, and also helps with scalability with large dataset. The other is the cold start problem, which refers to the algorithms performing poorly on newly added users as there's insufficient data for collaborative filtering to work properly. Most collaborative filtering algorithms suffer from the later problem, including matrix factorization.

In this paper, we propose a hybrid algorithm that addresses the above issues. The hybrid approach takes advantages of user-review information as well as user network structure by constructing a new objective function based on normal matrix factorization approach. We compare the proposed hybrid framework with other state-of-the-art methods, and show that it outperforms those methods. We further extensively evaluate the hybrid framework by running it on different settings.

The goal of the project is to predict user's rating on businesses based on users past experience, business history, and user's friendship network structure. Our final deliverable system is a rating predictor. The input of the system is organized into four part: (1) User, (2) User Friendship, (3) Business, and (4) Review. The output of the system is a list of user-business review rating.

Our contribution lies in following aspects:

- We propose to use random walk to mine the underlying structure of network structure and construct user similarity graph.
- We design a objective function which leverages both collaborative filtering and graph regularization, and develop a corresponding algorithm to optimize it.
- The experiments demonstrate that our proposed

[1]We will be using the terms rating predication and recommendation interchangeably through out this paper, since we can treat recommendation as a $0-1$ rating prediction problem.

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

method has best performance over several baselines and contributes to solve the cold start problem.

The rest of this paper is organized as follows: Section II specifies previous related work. Section III elaborates various algorithms we utilize to approach the problem including the proposed hybrid method. In section IV, we outline our evaluation method, present and analyze the experiment results. Section V summarizes our contributions to this problem. Section VI refers to some related work that could be done with more time and resources in the future. Section VII lists the contributions of each individual author.

II. Related Work

Rating prediction and recommendation of user reviews on Yelp is an important paradigm that discovers interesting business for users, helps business understand their service, therefore attracts a lot of attentions in recent years. In this section, we briefly introduce the related work in this area, by classifying them into three categories according the fundamental ideas of these approaches.

- ***Similarity-based recommendation***: This category of methods recommend a business to a user based on the attributes of user or business, e.g., description of the business, the profile of the users' interests, etc. Qu *et al.* utilize bag-of-opinions representation of users' reviews with a root word, a set of modifier words and negative words, and apply rigid regression to perform the review rating prediction [9]. Lee *et al.* applies support vector machines (SVMs) to the corporate credit rating problem in an attempt to suggest a new model using features of companies [10]. Kim *et al.* represent items with keyword features by a content-based filtering algorithm, and propose a community recommendation procedure for online readers [4].
- ***Collaborative filtering based recommendation***: This category of methods was successfully applied and wildly used in traditional recommendation and rating predication system, and is based on the assumption that similar users are likely to have similar ratings or like similar items. Rennie *et al.* investigate a direct gradient-based optimization method for maximum margin matrix factorization and demonstrate its' performance on large collaborative prediction problem [6]. Chen *et al.* propose an improved collaborative filtering method named combinational collaborative filtering (CCF), which considers multiple types of co-occurrences in social data and recommends personal communities [3]. Lawrence *et al.* develop a non-linear probabilistic matrix factorization using Gaussian process latent variable models to predict movie ratings on EachMovie. Yu *et al.* propose a collaborative filtering recommendation algorithm for Web communities, in which the latent links between communities and members are utilized to handle the sparsity problem [5].
- ***Hybrid recommendation***: This category of methods combines several recommendation method to recommend items or predict the ratings of items. [12] presents a hybrid approach which combines content-filtering techniques with those based on collaborative filtering to recommend TV programs to users. Chen *et al.* compare association rule mining (ARM) and latent dirichlet allocation (LDA) for the community recommendation, and find that LDA performs consistently better than ARM when recommending a list of more than 4 communities [2]. Zheng *et al.* propose a tensor decomposition model for recommendation, which measures the latent relations between users and items by considering both tags and users social relations [8]. Zheng *et al.* also propose an approach which combines the topic model and collaborative filtering, and this method is demonstrated to have better performance than traditional CF and negative matrix factorization [7]. Li *et al.* apply clustering techniques to the collaborative filtering framework to solve the cold start problem, which demonstrate superior results in case of cold start [13].

Our proposed approach in this paper is a hybrid one, where we take the advantage of both similarity-based recommendation methods and the collaborative filtering based recommendation methods to further improve the performance of recommendation. In the technology aspect, we extend the traditional matrix factorization, perform random walk algorithm to mine the underlying network structure, and propose a novel objective function in which both user-business ratings and user similarity graph are considered. Further, an element-wise optimization approach is proposed to solve the objective function and can be easily paralleled to speed up the method.

III. Algorithms and Model

In this section, we shows the detail of our hybrid approach based on user similarity graph from random work and matrix factorization. The main framework of our method is presented in Figure 1. The rest of this section is organized as follows, first we presents similarity based recommendation methods and how we generate the user similarity graph. Second, we give the basic the objective function of matrix factorization. Finally, we proposed our hybrid method based on network structure and collaborative filtering.

A. Similarity based recommendation

In this basic approach, we explored ways to calculate the similarity between users, which are then used to predict users' ratings on business and make recommendations. We compared two different ways to construct the user-user similarity matrix. The first approach is using random walk on user-user friendship data. The second approach is using traditional similarity heuristic on user review information.

Once a user-user similarity matrix has been created, user's ratings can be predicted based on the following equation.

$$r_{u,i} = k \sum_{v \in V} W(u,v) r_{v,i} \tag{1}$$

where u is the target user, $r_{u,i}$ is the rating that user u gave to business i, V is the set of users that rated business i, $W(u,v)$ is similarity calculated using different approaches, k is the normalizing factor that ensures $W(u,v)$ sums to one

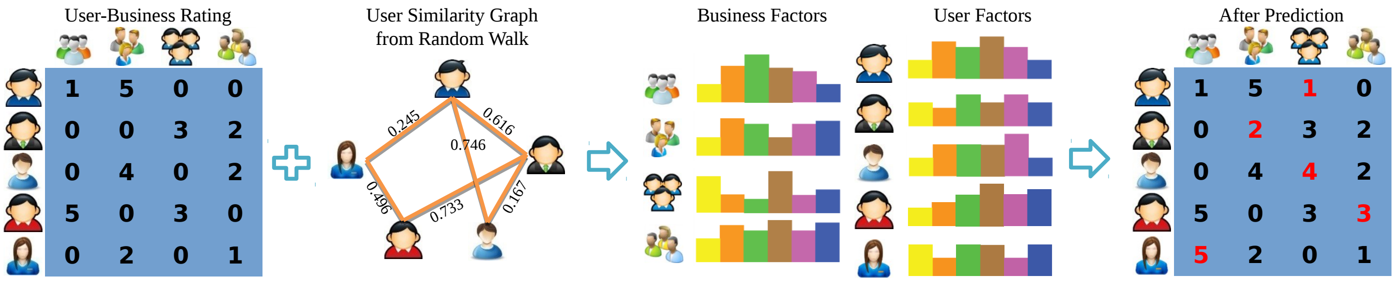


Fig. 1. The main framework of our proposed hybrid method. The user-business relationship and user similarity graph generated by random walk are integrated to obtain the user and business latent factor matrices. Specifically, the similarity graph is computed based on the features extracted from images and annotated tags. Then the recommendation results can be calculated by recovering the user-business relationship from the latent factors of users and businesses.

(i.e. $k = 1/\sum_{v \in V} W(u,v)$). The algorithm is based on the intuition that users would trust similar users' recommendations more. So higher weight is given to users with high similarity to target user. One major disadvantage of this similarity based recommendation is that if the similarity scores between the target user and users who reviewed business i are all zero, we would get zero predicted rating on business i.

Random walk with restart: In real life, we often find that our friends usually share similar tastes with us. For example, Chinese people, which are more likely to be friends with Chinese people, would have preference on authentic Chinese food to American-Chinese food. Students may prefer bars with more students.

Random walk is an effective algorithm on calculating the proximity of nodes in graph. We apply random walk with restart on user-user friendship graph $G(V,E)$ to estimate the similarity between users.

$$s_u = (1-\beta) \sum_{v \in V} \frac{s_v}{d_v} + \beta \mathbb{1}\{u = u^*\} \tag{2}$$

where u^* is the target user, s_u is the proximity(similarity) between user u and user u^*, β is the restart probability, V is the set of users connected to user u^*, d_v is the degree of user v, and $\mathbb{1}\{u = u^*\}$ is the indicator function. In our similarity based model, s_u would be the $W(u^*, u)$ in equation (1). The time complexity of random walk is $O(k|E|)$ where k is the number of iterations and $|E|$ is the number of edges in graph G. This is because in each iteration, the algorithm updates the proximity from neighborhoods nodes. Thus computing the whole similarity matrix would cost $O(k|E||V|)$. To improve computation efficiency, we can adopt the fast random walk algorithm proposed in [14].

Traditional similarity heuristic: Instead of using user friendship information, we could also take advantage of users' review data. Traditional similarity heuristic such as Pearson correlation and cosine similarity is a commonly used technique to compute the similarity between users. The rationale is that users that rated same businesses similarly are more likely to share similar tastes.

We chose to use Cosine similarity simply because it doesn't have the problem of Pearson correlation that users with few shared items in common would have very high similarities (e.g., one pair is undefined, two pairs is 1). The Cosine similarity is given by the following equation,

$$W(u,v) = \frac{\sum_{i \in I_{uv}} r_{u,i} r_{v,i}}{\sqrt{\sum_{i \in I_u} r_{u,i}^2} \sqrt{\sum_{i \in I_v} r_{v,i}^2}} \tag{3}$$

where I_{uv} is the set of businesses that user u and user v both rated, I_u is the set of business that user u rated, and $r_{u,i}$ is the rating that user u gave to business i. The time complexity to compute the whole user similarity matrix is $O(rn)$, where r is the number of reviews, and n is the number of users. This is because after pre-computation, the algorithm takes $O(\frac{r}{n})$ to calculate equation (3), and we need to compute it $O(n^2)$ times.

B. *Collaborative filtering*

We construct a matrix R, with users as the rows and businesses as the column. We then design a matrix factorization based method to predict the rating of a business for a specific user based on the latent factor we learned from the matrix factorization.

Given the sets of M users, N businesses respectively, $R \in \mathbb{R}^{M \times L}$ is the affiliation matrix between users and businesses, where $R_{ij} = 0$ means that the ith user did not review the jth business, $R_{ij} = s$ means that the ith user reviewed the jth business with a rating of s. Given the information above, the aims of review recommendation is to recover a new affiliation matrix R_{rec} denote the relationship between users and businesses, and more importantly predict ratings of users based on R_{rec}.

The traditional matrix factorization approach for recommendation tries to factorize the affiliation matrix R into two $M \times K$ and $N \times K$ dimensional low-rank matrices U and G by:

$$\underset{U,G}{\operatorname{argmin}} \parallel R - UG^T \parallel_F + \lambda(\parallel U \parallel_F + \parallel G \parallel_F) \tag{4}$$

where $\| \cdot \|_F$ denotes the Frobenius norm and K is dimensionality of the latent factors of both users and businesses. Besides, the regularization penalties $\lambda(\parallel U \parallel_F + \parallel G \parallel_F)$ is utilized in order to avoid over-fitting. Afterwards, the recommendation results can be obtained by calculating similarity between the latent factors of users and businesses as $R_{rec} = UG^T$.

C. Hybrid Approach Based on Network Structure and Collaborative Filtering

As in [15], recommendation suffers from the cold start problem. Also, many recommendation techniques fail to take the network structure into consideration. To solve these problems, we need to fully exploit the user similarity information and the network structure information. Therefore, we proposed to combine the traditional collaborative filtering approach with random walk based network structure mining approach.

Similar to other collaborative filtering algorithms, matrix factorization based recommendation still suffers *cold-start* problem, i.e., recommendation results for new users who have not reviewed any business tend to be very inaccurate. This problem could be solved, to some extent, by exploiting network information of the users, i.e., the underlying network structure and the neighborhood information of users. The basic assumption is: if two users have similar network structure or in the similar neighborhood of the graph, they tend to have similar rating the businesses. Based on this assumption, we can incorporate such user similarity graph W constructed from random walk into the matrix factorization based rating prediction framework.

To guarantee that users who have higher similarity in network structure also obtain similar latent factors in matrix factorization, we introduce the following graph regularization term:

$$\begin{aligned}
&\frac{1}{2}\sum_{i,j} \parallel u_i - u_j \parallel_2 W_{ij} \\
=&\sum_{ij} u_i W_{ij} u_i^T - \sum_{ij} u_j W_{ij} u_j^T \\
=&\sum_{i} u_i D_{ii} u_i^T - \sum_{ij} u_j W_{ij} u_j^T \\
=&tr(U^T(D-W)U) \\
=&tr(U^T L U),
\end{aligned} \tag{5}$$

where $tr(\cdot)$ denotes the matrix trace, $D_{ii} = \sum_j W_{ij}$ is a diagonal matrix and $L = D - W$ is the Laplacian matrix of the user similarity graph.

By leveraging the collaborating information and user similarity graph, we propose our hybrid method, which unifies matrix factorization and graph regularization as:

$$\underset{U,G}{\operatorname{argmin}} \ \parallel (R - UG^T) \parallel + \mu tr(U^T L U) + \lambda(\parallel U \parallel_F + \parallel G \parallel_F) \tag{6}$$

where μ is the trade-off parameter between collaborating information and user similarity graph, L is the Laplacian matrix of the network structure based user similarity graph. $\lambda(\parallel U \parallel_F + \parallel G \parallel_F)$ is the regularization term to avoid overfitting.

D. Optimization

Here, we describe the optimization of our hybrid method. Although the proposed objective function is not a convex function of U and G, but it is convex to one variable when the other one is fixed. Therefore, we could obtain the local optimal solution by alternatively updating the two variables using gradient descent methods.

```
Data: matrix R, matrix W, an error threshold ϵ
Result: U, G
initialization U ∈ R^{M×K}, G ∈ R^{N×K} with small
random values;
set η as step size;
D_ii = Σ_i W_ij;
L = D − W;
while Loss_t − Loss_{t+1} > ϵ do
    forall the R_ij > 0 do
        Y_ij = U_i* × G^T_j*;
        U_i* ← U_i* − (Y_ij − R_ij)Gj* − 2μL_i × U − 2λU_i;
        G_j* ← G_j* − (Y_ij − R_ij)Ui* − 2λG_j
    end
end
return U, G
```

Algorithm 1: Algorithm of our proposed hybrid method

Denoting the objective function as $J(U, G)$, we can calculate the gradient of $J(U, G)$. The partial derivative with respect to U is:

$$\frac{\partial J}{\partial U} = (UG^T - R)G + 2\mu LU + 2\lambda U \tag{7}$$

The partial derivative with respect to G is:

$$\frac{\partial J}{\partial G} = (UG^T - R)^T U + 2\lambda G \tag{8}$$

In our model, W influences factor matrix R. The dense representation of W contributes to the generation of a relative accurate U, which reduces the decomposition error of R_{rec} in turn. In other word, the user similarity from the underlying network structure is propagated into recovered matrix R_{rec}

Algorithm 1 presents the algorithm for the collaborative matrix factorization. As there is no closed-form solution for finding the global optimal result of the objective function, we used a numeric method, gradient descent, to find a local optimization. Specifically, we use an element-wise optimization algorithm, which updates each entry in the matrix independently and can be easily parallelled.

IV. Evaluation and Result

In this section, we evaluated different algorithms on real-world data from Yelp to compare the effectiveness and efficiency of them.

A. Data Set

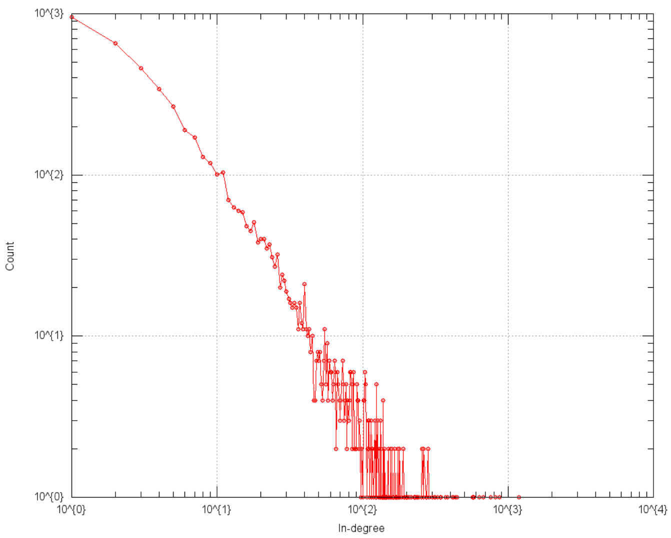
To evaluate the performance of the proposed approaches, we use user-business review dataset from Yelp Dataset Challenge [1]. The dataset contains users' basic information including friendship, business information, and review data (rating from 1 to 5, time, etc.). We are specifically interested in the information that constructs network such as friendship and reviews. Our approaches don't take advantages of node features as that usually requires domain knowledge and complex feature engineering.

Since business recommendation is mostly location-based, we only selected the businesses from Phoenix to experiment with. We further eliminated businesses with few reviews (less

TABLE I. DATASET STATISTICS

Number of Users	5703
Number of Businesses	5224
Number of Reviews	99581
Average Review Rating (1-5)	3.82
Number of Nodes (User Graph)	5703
Number of Edges (User Graph)	42493
Connected Component Size (User Graph)	0.84
Number of Connected Components (User Graph)	276
Alpha of Power-law (User Graph)	1.58

Fig. 2. Power-law Distribution of Users in the Yelp Dataset



than 5), and users with few friends (less than 5), and only kept reviews between these businesses and users. After the preprocessing, we have our dataset statistics presented in Table I. Fig. 2 shows the power-law distribution of the user graph.

To evaluate the performance of rating prediction, we randomly sample $\alpha \times 100\%$ of the user-business ratings from the user-business affiliation matrix to generate the matrix R for training and use the remaining $(1-\alpha) \times 100$ review data as the ground-truth for evaluation.

B. Evaluation Metrics

To comprehensively evaluate the performance of the proposed approach, we consider the evaluation metrics Root Mean Squared Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}} \tag{9}$$

where $\hat{y}_t$ is the predicted value, y_t is the ground truth value and n is the number of predictions.

C. Recommendation Performance Comparison

In order to demonstrate the effectiveness of the proposed approach, we implement the following approaches and compare the performances:

1) **Random Rating**: Randomly predict ratings from 1 to 5.

TABLE II. RATING PREDICATION PERFORMANCE COMPARISON ($\alpha = 70\%$)

Method	RMSE	Runtime (min)
Random Rating	1.73	**0**
Cosine Similarity	1.42	1
Random Walk ($\beta = 0.2$)	1.12	59
Random Walk ($\beta = 0.1$)	1.11	65
Matrix Factorization	1.15	72
Hybrid method	**1.02**	162

2) **Random Walk**: Recommendation based on similarity matrix generated by random walk on user friendship graph.
3) **Cosine Similarity**: Recommendation based on similarity matrix generated by computing Cosine Similarity on user-review information.
4) **Matrix Factorization**: Matrix factorization based recommendation by using user-business relationship.
5) **Hybrid method**: Our proposed hybrid method that leverages between newtork structure and matrix factorization by using user-business relationship.

Parameter Settings In this experiment, parameter μ is set empirically to 0.1 and λ is set to 0.1. The dimensionality of latent factors K is set to 10. We also compared two different restart probabilities ($\beta = 0.2$ and $\beta = 0.1$) on the Random Walk approach. We fed the similarity matrix generated by Random Walk with $\beta = 0.1$ into our hybrid approach. It should be noted that the parameters of all the competitive methods have been fairly tuned using cross-validate, and the average evaluation results after 10-fold cross-validation are selected.

D. Results and Analysis

Table II shows the performance comparison of above rating prediction approaches when $\alpha = 70\%$ in terms of evaluation metrics RMSE and runtime. The runtime was collected from running algorithms on a Macbook Pro with 4-core 2.3GHz CPU and 16GB memory. The results in bold indicate the best performance. From Table II, it can be observed that our proposed hybrid method largely outperforms the other state-of-the-art approaches in terms of RMSE. The superior performance of the proposed approach comes from two aspects, one is the selection of matrix factorization, while the other one is the integration of user similarity graph learned from random walk algorithm. One thing to note about the two similarity based approaches is that, as mentioned before, when the target user has zero similarities with all the other users who reviewed business i, the prediction would be zero. In such scenario, we evenly distributed the weight on all the users who reviewed business i in order to make a rather reasonable prediction. In the Cosine Similarity method, we've encountered 1849 out of 30099 such cases, which is a major reason why the performance is much worse than other methods. Compared with Cosine Similarity, Random Walk has only 59 out of 30099 special cases. The reason behind this is mainly that Cosine Similarity matrix is much sparser than Random Walk matrix due to the fact that Cosine Similarity is zero unless two users have rated some businesses in common while Random Walk probability is nonzero unless a node is disconnected from the target node or the probability is too small to be represented. Once we took out those special cases from our test set, the RMSE of Cosine Similarity approach became close to Random

Walk method (RMSE(Cosine Similarity) = 1.11). However, we shouldn't compare those two methods under such metric since though the RMSEs are similar, we would be able to predict and make recommendation on less user-business pair by using Cosine Similarity compared with Random Walk. Furthermore, as we've noticed, Random Walk with smaller β (restart probability) performs better. This corresponds to the fact that smaller β allows the walker to explore more of the network structure, which makes the random walk probability reflects the real proximity better.

On the runtime side, our hybrid method took the longest time to run. However, it's worth noting that since this method requires the construction of the similarity matrix, we've included the construction time (which in this case is 65 min) in the final runtime. Additionally, our proposed element-wise matrix optimization method for the hybrid method optimizes a nonzero element at a time, therefore can be easily paralleled. The major advantage of Cosine Similarity matrix is it runs much faster (almost two magnitudes) than other algorithms. It's also worth mentioning that Random Walk with smaller β runs slightly slower than that with larger β since it takes more runs to converge.

E. Performance Evaluation of the Hybrid Method

We conducted further experiments on the hybrid method to fully evaluate its performance. First, to understand whether our hybrid method resolves the cold start problem, we compared the hybrid framework and matrix factorization method side by side on different sizes of training set. Figure 3 shows the rating prediction performance comparison with the proportion of training data α varies from 10% to 90% in terms of RMSE. The superior performance at different α further verified effectiveness of the proposed approach. It can be observed that the accuracy of hybrid method slightly increases as α gets larger. Also, when α is small, which means we have few historic user review data (i.e. cold start problem), the hybrid method still performs reasonably, while performance of matrix factorization method degrades drastically. For example, at $\alpha = 0.1$, we only trained on 9958 reviews (average number of reviews per user is 1.75), where RMSE of hybrid approach achieves 1.29 and RMSE of matrix factorization is only 2.31. The reason is that the user similarity regularization plays a important role when the initial review rating matrix is sparse, which is consistent with our motivation. This proves that integrating user similarity graph generated by random walk into collaborative filtering can help solve the cold start problem.

In the proposed hybrid method, we utilize both user-business rating matrix R and user similarity graph W from network structure. In the above experiment, we factorize R into user factor matrix U and business factor matrix G. U and G are low-rank matrices that $U \in \mathcal{R}^{M\times K}$ and $G \in \mathcal{R}^{N\times K}$. In this section, we evaluate the performance of the hybrid method using various K, and evaluate the impact of parameter K to performance of the hybrid method in RMSE.

Figure 4 shows the performance of the hybrid method and the impact of K to it. In previous experiments, we use $K = 10$. From Figure 4, it could be observed that the increase of K improves the performance of recommendation in terms of RMSE, which makes sense as the factor matrices have more

Fig. 3. Performance of our hybrid method with α varies from 10% to 90%

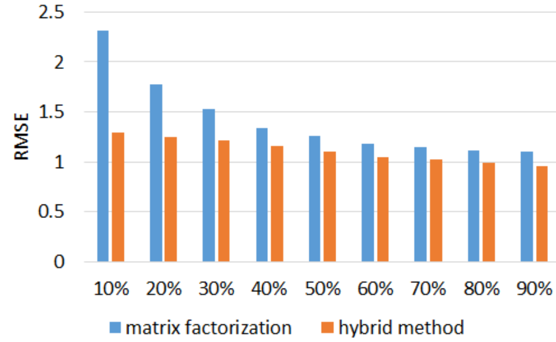
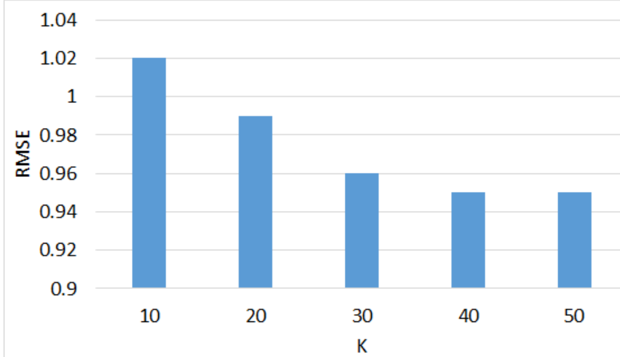


Fig. 4. Performance of our hybrid method with K varies from 10 to 50



dimensions that can better recover the user-business rating matrix. Further, as K continues to increase, the performance of recommendation tends to be stable with RMSE $= 0.95$.

V. Conclusion

In this paper, we proposed a hybrid framework for rating prediction based on user network structure properties and past review information. We first constructed a user similarity graph using random walk of user network structure. Then, we designed an objective function to optimize based on matrix factorization and user similarity graph regularization. We developed an element-wise optimization based on gradient descent methods. Experiments on Yelp review dataset demonstrate that our method outperforms other state-of-the-art algorithms including similarity based recommendation methods and collaboretive filtering using matrix factorization. Besides rating prediction, our framework can be easily adapted to item (e.g. business) recommendation as well.

VI. Future Work

We plan to change our element-wise matrix factorization algorithm to parallel in order to improve the scalability of our algorithm.

Also, we hope to use more additional data sources (e.g. user check in data, social media data) and incorporate them into our hybrid recommendation method to further improve the performance.

VII. CONTRIBUTORS

We are group 45, the contributions of individual team members are listed as below:

- Yilun Wang: report writing, implementing matrix factorization and the hybrid method, tabulating and plotting final results
- Guoxing Li: report writing, data pre-processing, implementing random walk, cosine similarity based recommendation
- Jingrui Zhang: report writing, writing code to get basic statistics of data, running tests

ACKNOWLEDGMENT

The authors would like to thank Prof. Leskovec and TAs for useful discussion. Thanks Vikesh Khanna for the useful advise and following feedback about this project.

REFERENCES

[1] Yelp Dataset Challenge. http://www.yelp.com/dataset_challenge

[2] Chen, W.Y., Chu, J.C., Luan, J., Bai, H., Wang, Y., Chang, E.Y.: Collaborative filtering for orkut communities: Discovery of user latent behavior. In: Proceedings of the 18th international conference on World wide web. pp. 681–690 (2009)

[3] Chen, W.Y., Zhang, D., Chang, E.Y.: Combinational collaborative filtering for personalized community recommendation. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 115–123 (2008)

[4] Kim, H.K., Oh, H.Y., Gu, J.C., Kim, J.K.: Commenders: A recommendation procedure for online book communities. Electronic Commerce Research and Applications 10(5), 501–509 (2011)

[5] Qian, Y., Zhiyong, P., Liang, H., Ming, Y., Dawen, J.: A latent topic based collaborative filtering recommendation algorithm for web communities. In: Proceedings of the Ninth Web Information Systems and Applications Conference. pp. 241–246 (2012)

[6] Rennie, J.D.M., Srebro, N.: Fast maximum margin matrix factorization for collaborative prediction. In: ICML. pp. 713–719 (2005)

[7] Zheng, N., Bao, H.: Flickr group recommendation based on user-generated tags and social relations via topic model. In: 10th International Symposium on Neural Networks. pp. 514–523 (2013)

[8] Zheng, N., Li, Q., Liao, S., Zhang, L.: Flickr group recommendation based on tensor decomposition. In: Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. pp. 737–738 (2010)

[9] Qu, L., Ifrim, G., & Weikum, G. (2010, August). The bag-of-opinions method for review rating prediction from sparse text patterns. In Proceedings of the 23rd International Conference on Computational Linguistics (pp. 913-921). Association for Computational Linguistics.

[10] Lee, Y. C. (2007). Application of support vector machines to corporate credit rating prediction. Expert Systems with Applications, 33(1), 67-74.

[11] Lawrence, N. D., & Urtasun, R. (2009, June). Non-linear matrix factorization with Gaussian processes. In Proceedings of the 26th Annual International Conference on Machine Learning (pp. 601-608). ACM.

[12] Barragns-Martnez, A. B., Costa-Montenegro, E., Burguillo, J. C., Rey-Lpez, M., Mikic-Fonte, F. A., & Peleteiro, A. (2010). A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition. Information Sciences, 180(22), 4290-4311.

[13] Li, Q., & Kim, B. M. (2003, October). Clustering approach for hybrid recommender system. In Web Intelligence, 2003. WI 2003. Proceedings. IEEE/WIC International Conference on (pp. 33-38). IEEE.

[14] Tong, H., Faloutsos, C., & Pan, J. Y. (2008). Random walk with restart: fast solutions and applications. Knowledge and Information Systems, 14(3), 327-346.

[15] Khrouf, H., & Troncy, R. (2013, October). Hybrid event recommendation using linked data and user diversity. In Proceedings of the 7th ACM conference on Recommender systems (pp. 185-192). ACM.