FRFP: A friend recommendation method based on fine-grained preference*

Mingmin Shao[1], Wenjun Jiang[2], and Lei Zhang[3]

[1] Hunan University, Changsha, China
shaomingmin@hnu.edu.cn
[2] Hunan University, Changsha, China
jiangwenjun@hnu.edu.cn
https://sites.google.com/site/happywenjunjiang2012/
[3] Hunan University, Changsha, China
zhang_lei@hnu.edu.cn

Abstract. With the development of social networks, the traditional photography community has become a social platform for photography enthusiasts. In photography community, users are often asked and encouraged to give relevant tags based on the content of the photos when uploading them. These tags are often fine-grained and can be better used to analyze the user's fine-grained photography preferences for friend recommendation. However, recommendation is challenging, since state-of-the-art method does not pay attention to the user's photography preferences is fine-grained, resulting in poor friend recommendation. Therefore, we try to propose a new Friend Recommendation method by user's Fine-grained Preference (FRFP). Firstly, FRFP method extracts the user's fine-grained photography preference features from the perspective of the fine-grained tag. Then, we use the pagerank algorithm to calculate the importance of the preference feature tag as the score of the user-item scoring matrix, and generate a friend recommendation list through the collaborative filtering algorithm. Finally, we use user activity to weight the users in friend recommendation list, preferentially recommend users with high user activity to target user, and improve the quality of friend recommendation. The experimental results on real-word data show the effectiveness and precision of the proposed method in friend recommendation for photographers.

Keywords: Photography community · Fine-grained preference · Social network · Recommended friends.

1 Introduction

1.1 Research background

In the photography community, photographer can share their photos, communicate and interact with other users, and make friends who have similar photography preferences. For the current popular photography communities (such as

* Supported by organization x.

flickr) , one challenging issue is how to help these users to efficiently find new social friends. They cannot be able to provide a method to accurately recommend potential friends for photographers with similar photography preferences.

In the photography community, users are encouraged to give relevant tags based on the content of their photos. Therefore, those conditions provide the possibility to analyze the user's fine-grained photography preferences and improving the accuracy of recommend potential friends to users with similar fine-grained photography preferences.

However, in the recommendation of friends in the photography community, there are the following problems: (1) Users in the photography community have specific photography styles and hobbies, and traditional coarse-grained friend-recommendation methods cannot meet user needs. (2) The number of tags for user photos tends to be large and there is no suitable method for processing the tag data to obtain the user's fine-grained photography preferences. (3) The existing friend-recommendation method has low accuracy and quality, which has little significance in practice.

We are inspired by the recent work by calculating similarity from users interest topics to find potential friends with the same interest in social tagging systems(Buxiao Wu *et al.*, 2015). Because the current photography community contains a huge number of tags, it is a very complicated tag system. However, for the photography community, helping users find friends with similar fine-grained photography preferences can increase user satisfaction and increase their competitiveness in the industry. Therefore, all the work in this paper is based on the following three motivations: (1) Design a suitable method to extract the user's fine-grained photography preferences in the tag system. (2) Calculate the similarity of photography preferences between users, and improve the accuracy of friend-recommendation. (3) The more active users are preferentially recommended when the user preference similarity is close, so that improve the quality of the friend recommendation.

In summary, this paper makes the following contributions:

(1) We propose a friend recommendation method called FRFP to obtain the user's fine-grained photography preferences through tags.

(2) We have conducted comprehensive experiments to show that the proposed method significantly improves the accuracy of friend-recommendation.

(3) We use user activity to improve the quality of friend recommendations.

The rest of this paper is organized as follows: Section 2 will discuss related work. Section 3 and Section 4 introduces our method and potential friend recommendation. Section 5 shows the analysis and comparison of experimental results. Section 6 summarize this paper and discuss future work.

2 Related Work

Researches about friend recommendation related to our work and divided into two fields: (1) Friend recommendation based on user connections in social networks. (2) Friend recommendation based on similarity of interest preference.

2.1 Friend recommendation based on user connections in social networks

In social networks, the connection of users can be seen as a large directed graph. According to the classic Triadic-Closure theory [8] , we can recommend potential friends to target user by analyzing the connections of all users and counting the number of common friends in the social network. For example, QQ, Sina Weibo, etc.

Based on the user's connection, the friend recommendation problem can be roughly classified into three categories according to the existing algorithms [9] [16]. The classification method mainly uses the features between two user nodes, such as the path-based metric Katz or the neighbor-based metric Adamic/Adar. Predicting the value of friendship by using the SVM model to train the two-classifier for the friendship feature state between users [7]. The fitting method represents the friendship between users and as close as possible to the observed friendship value, for example, using Matrix Factorization to predict the value of an unknown friendship [12] [17]. Due to the sparseness of data in open social network, the classification and fitting methods cannot handle the serious data imbalance problems [4]. Thus, the common Bayesian personalized ranking model [21] [22] is offen used to solve the imbalance problem effectively.

Huanyang Zheng *et al.* [25] study the friend recommendation strategy from the perspective of maximizing social influence. For system providers (such as Facebook), the goal is to recommend a fixed number of new friends, so that the user can maximize his/her social influence by making new friends. Shangrong Huang *et al.* [13] Align the user's social relationship network and tag network to generate a possible friend list.

2.2 Friend recommendation based on similarity of interest preference

J. Hannon *et al.* [10] considers user generated content on Twitter and develops a fast algorithm for real-time user-to-user similarity for follower/followee recommendation. It is developed only for text similarity. Xie Xing *et al.* [24] designed an ordinary friend recommendation framework, which can represent user interest in two aspects: context (location, time) and content, and combine domain knowledge to improve recommendation quality.

Shangrong Huang *et al.* [13] conduct the three-way clustering by users, tags and images to obtain topics that might interest user. Re-recommended the friend recommendation list generated by aligning the user network and the tag network to optimize the recommendation effect. Mohammed Mehedi Hasan *et al.* [11] analyzed the diversification of users in the social network and the dynamism of user interests, measured the frequency of activities completed by users and updated the data set according to the activity. Buxiao Wu *et al.* [23] proposed a Friend Recommendation algorithm by User similarity Graph (FRUG) to find potential friends with the same interest in social tagging systems.

3 Problem definition and preliminary concepts

In this section, we formulate the problem we address, and provide some preliminary concepts. In the photography community, each user will be given relevant tags based on the content of the photos. However, not all tags are useful and can directly express the user's fine-grained photography preferences, so a large number of tags need to be pre-processed to get tags that can represent fine-grained photography preferences.

The FRFP method proposed in this paper processes a large amount of tags to obtain a few tags that can represent the user's fine-grained photography preferences. The FRFP method mainly includes the following steps (Fig. 1): (1) Performing vectorized representation of user's tags. (2) Using cosine similarity to calculate similarity between word vectors. (3) Clustering tags based on the similarity of word vectors. (4) Extracting the high frequency tags and the tags close to each cluster center in same cluster, and use these tags representing the user's fine-grained preference features.

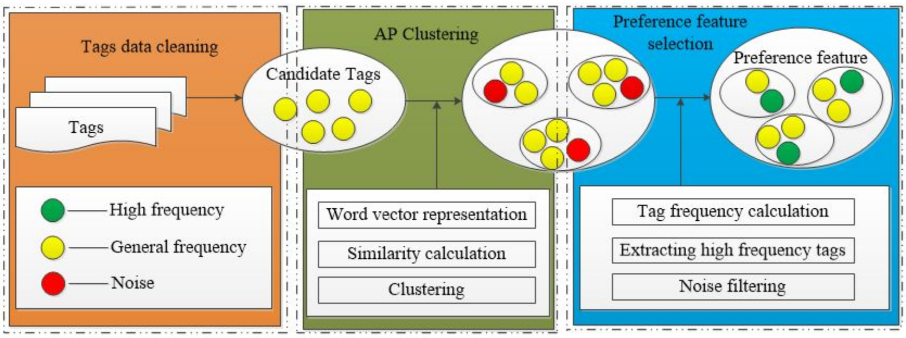

Fig. 1. Fine-grained preference feature selection.

3.1 Problem definition

Given a set of metadata (U, T, I), U is a set of users, T is a set of user's fine-grained tags and I is a set of user's information. For two different users, u and v in U, we seek to determine how to design an efficient method to extract user's fine-grained preference features, and caculate preference similarity of u and v. Specifically, how to recommend friends such a v with similar preference for u , and how to combine user's information to improve the quality of friend recommendation simultaneously. Our goal is to design an effective method to extract user's fine-grained preference features from tags, improve the precision and quality of friend recommendation based on user's information.

3.2 Preliminary concepts

In social network-based recommendation system, friend recommendation is the essential part. In this paper, we first extract the user's fine-grained preference features. Secondly, the friend list is generated based on users' fine-grained preference similarity. Finally, users' similarity are weighted to generate a final friend recommendation list. The three key concepts of FRFP are as follows:

Definition 1. (H-tag). In the original tagging system, we first remove the noise tags and get a collection of the candidate tags. Then, we count the frequency of the candidate tags and the tag with a frequency greater than the average frequency is defined as H-tag.

Definition 2. (C-tag). After clustering the candidate tags, we define a tag with Euclidean distance to the cluster center less than the average distance of the rest as C-tag.

Definition 3. (A_u). $A_u \in (0, 1)$ represents the activity of the user u, and the greater number of photos, fans, and pageviews, the greater activity of user will be. When the users' fine-grained preference similarity is close, the user with high activity is preferentially recommended as the friend of the target user.

3.3 Solution overview

In the photography community, each user will be given a relevant tag based on the content of the photo. But not all tags are useful and can directly express the user's fine-grained photography preferences, so a large number of tags need to be pre-processed to get tags that can represent fine-grained photography preferences. The FRFP method proposed in this paper processes a large amount of tag data to obtain a tag that can represent the user's fine-grained photography preferences. The FRFP method mainly includes four tasks are as follows:

Task 1: Word vector representation of fine-grained tags. We uses the skip-gram model in the word2vec [18] tool, through the Wikipedia article [1] provided by Matt Mahoney as the English corpus(containing 1 billion characters) for model training. Distribution-representation [19] (200-dimensional word embedding) method is used to obtain the similarity between two words, because it effectively reflect the semantic relationship between two words than traditional one-hot representation. The skip-gram [15] model includes steps are as follows:

Given a sequence of training words $w_1, w_2, ..., w_T$, the objective function of the word vector model is to maximize the average log probability.

$$\frac{1}{T}\sum_{t=k}^{T-k} \log p(w_t|w_{t-k}, \cdots, w_{t+k}) \qquad (1)$$

where T is the number of training words.

The prediction task is typically done via a multiclass classifier, such as softmax. There, we have

$$p(w_t|w_{t-k}, \cdots, w_{t+k} = \frac{e^{y_{w_t}}}{\sum_i e^{y_i}}) \qquad (2)$$

Each of y_i is un-normalized log-probability for each output word i, computed as

$$y = \psi + \vartheta\varphi(w_t|w_{t-k}, \cdots, w_{t+k}; W) \qquad (3)$$

Where ϑ, ψ are the softmax parameters, W is a word vector matrix, and φ is constructed by a concatenation or average of word vectors extracted from W.

Task 2: Calculate similarity between word vectors. After word vectorization representation of fine-grained tag, we obtain the similarity between fine-grained tags by calculating the cosine similarity between word vectors. We have

$$\cos(w_i, w_j) = \frac{\sum_{k=1}^{N}(W_i^k \times W_j^k)}{\sqrt{\sum\limits_{k=1}^{N}(W_i^k)^2} \times \sqrt{\sum\limits_{k=1}^{N}(W_j^k)^2}} \qquad (4)$$

Where w_i^k represents the k_{th} dimension of the word vector of the candidate fine-grained tag w_i, N represents the dimension of the vector, w_j^k represents the k_{th} dimension of the word embedding of the candidate fine-grained tags w_j.

Task 3: Tags clustering based on similarity between word vectors. In this paper, the Affinity Propagation (AP) clustering [3] method is used to cluster tags based on similarities of word vectors. The responsibility information and the availability information of similarity data is continuously iteratively updated, until a stable cluster center is generated. The steps of AP clustering algorithm as follows:

Update the responsibility information formula $r_{t+1}(m,n)$ of the $(t+1)_{th}$ iteration.

$$\begin{cases} S(m,n) - \max\limits_{l \neq n}\{a_t(m,l) + r_t(m,l)\}, m \neq n \\ S(m,n) - \max\limits_{l \neq n} S(m,l), m = n \end{cases} \qquad (5)$$

Update the availability information formula $a_{t+1}(m,n)$ of the $(t+1)_{th}$ iteration.

$$\begin{cases} \min\{0, r_{t+1}(n,n) + \sum\limits_{l \neq m,n} \max\{r_{t+1}(l,n), 0\}\}, m \neq n \\ \sum\limits_{l \neq n} \max\{r_{t+1}(l,n), 0\}, m = n \end{cases} \qquad (6)$$

Where r represents responsility information, a represents availabilty information, S is a similarity matrix, and $S(m,n)$ takes a negative Euclidean distance of m and n. When $m = n$, $S(m,n)$ takes the minimum or median of the entire matrix. The larger value of the $S(m,n)$, the greater number of clusters that will eventually be produced. m and n respectively represent two data objects in the same cluster, t represents the number of iterations, $r_{t+1}(m,n)$ represents the responsility information of the $(t+1)_{th}$ iteration, and $a_{t+1}(m,n)$ represents the availabilty information of the $(t+1)_{th}$ iteration.

Task 4: Extract user's fine-grained preference features. In the photography community, H-tag(*Definition 1.*) and C-tag(*Definition 2.*) often reflect the user's fine-grained photography preferences. A plurality of clusters are formed after clustering candidate tags, and every cluster can represents one aspect of the user's fine-grained photography preferences. First, we delete the unqualified clusters (the number of tags is less than 3), and then select the cluster center tag, H-tag and C-tags in same cluster. Finally, the three kinds of tags are defined as user's fine-grained preference feature.

4 Potential friend recommendation

Firstly, we establishe a user-item scoring matrix by processesing the metadata of users and tags. Then, we use the collaborative filtering [5] algorithm to calculate the similarity of fine-grained photography preferences between the target user and other users, and generates a potential friend recommendation list. Finally, we use user activity to weight the users in friend recommendation list, preferentially recommend users with high user activity to target user, and improve the quality of friend recommendation (Fig. 2).

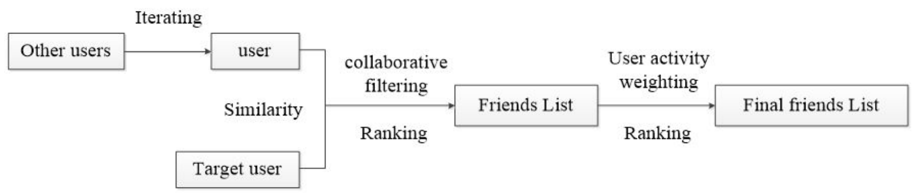

Fig. 2. Potential friend recommendation.

4.1 Calculate user similarity based on preference features

We use the R to assess the importance of the user's fine-grained photography preference feature tags. The tag frequency and similarity are greater, the larger R will be. In the collaborative-filtering algorithm, we use the value of R as a score in the user-item rating matrix.

We using the PageRank algorithm [20] to calculate the R value of preference feature tags, by the tag frequency and similarity. We define the fine-grained tag i of user u has a R value R_i^u, and all R values satisfy $R_i^u \in (0, 1)$.

In this work, we use jaccard similarity coefficient to evaluate the similarity between user u and v. $J(F^u, F^v)$ is defined as follows:

$$J(F^u, F^v) = \frac{|F^u \cap F^v|}{|F^u \cup F^v|} \qquad (7)$$

where F^u and F^v are N-dimensional vectors consisting of 0 and 1, N represents the number of different tags for all users. For example, $F^u = (1, 0, 1, ..., 0, 1)$, 1 means that the tag is included, and 0 means that the tag is not included. In the friend recommendation, we use the jaccard similarity coefficient to calculate the similarity of the preference characteristics of the target user and other users. We sort the users in descending order according to the similarity and generate a list of friend recommendations.

4.2 Filtering friends recommendation list based on user activity

In photography community, we need to use user's information to extract more attractive features, and these features often play an important role in friend recommendations. In this paper, we select the number of photos, fans and pageviews as features to assess the user's activity.

We define $\lambda_1, \lambda_2, \lambda_3$ as the weight of photos, fans and pageviews. We user entropy weight method (EWM) to determine the value of $\lambda_1, \lambda_2, \lambda_3$. Entropy concept has been widely employed in social and physical sciences. A mathematical theory of communication was proposed by Shannon. Entropy evaluates the expected information content of a certain message. Entropy concept in information theory can be considered as a criterion for the degree of uncertainty represented by a discrete probability distribution. Entropy idea can be effectively employed in the process of decision making, because it measures existent contrasts between sets of data and clarifies the average intrinsic information transferred to decision maker. To determine objective weight through Shannon entropy [6], the following procedure should be adopted.

Step 1: Normalization of the arrays of decision matrix (performance indices) to obtain the project outcomes p_{ij}:

$$Y_{ij} = \frac{X_{ij} - \min(X_i)}{\max(X_i) - \min(X_i)} \tag{8}$$

$$p_{ij} = Y_{ij} / \sum_{i=1}^{N} Y_{ij} \tag{9}$$

Step 2: Computation of the entropy measure of project outcomes using the following equation:

$$E_j = -\ln(n)^{-1} \sum_{i=1}^{N} p_{ij} \ln p_{ij} \tag{10}$$

Step 3: Defining the objective weight based on the entropy concept:

$$\lambda_i = \frac{1 - E_i}{k - \sum E_i} \tag{11}$$

We define A_u as the activity of the user u, and the greater the number of photos, fans, and pageviews, the greater the activity will be. When the users' fine-grained preference similarities are close, the user with high activity is preferentially recommended as friend of the target user. We normalize the number of photos, fans and pageviews of all users. $A_u \in (0,1)$ is defined as follows:

$$A_u = \lambda_1 \times \frac{w_u}{w_m} + \lambda_2 \times \frac{f_u}{f_m} + \lambda_3 \times \frac{p_u}{p_m} \qquad (12)$$

Where w_u, f_u and p_u are the number of photos, fans and pageviews respectively of user u. w_m, f_m and p_m are the maximum number of photos, fans and pageviews respectively. In which, $\lambda_1 + \lambda_2 + \lambda_3 = 1$, $\lambda_1 \in (0,1)$, $\lambda_2 \in (0,1)$, $\lambda_3 \in (0,1)$.

The preference feature similarities of potential friend recommendation list are multiplied by the user activity, and the result are defined as Rec-coefficient. We sort Rec-coefficient in descending order according to the size, and the top-k users are selected as the final friend recommendation list.

5 Experimental Analysis

5.1 Dataset

The dataset of this paper is mainly from the data of one of the world's largest photography communities, Flickr [2]. The dataset mainly including tags and user's information such as id, the number of pageviews, fans and photos. In this work, we focus on analyzing the 2170 users' fine-grained preferences features through tags and recommending potential friends with similar fine-grained preferences for the target user.

Table 1. Dataset Statistics.

User	2,170
Photos	17,087
Tags	53,438 tags from 17,087 photos

Tagging systems can provide users effective ways to collaboratively annotate and organize items with their own tags. However, the flexibility of annotation brings with large numbers of redundant tags. It is a very difficult task to find users interest exactly and recommend proper friends to users in social tagging systems. Therefore, we need to pre-process the tag data to get the candidate tag set (Table 2): (1) Delete all the fine-grained tags whose frequency of the tag is less than 2. (2) Delete fine-grained tags with misspellings. (3) Converts all uppercase letters to lowercase letters. (4) Delete tags that cannot be vectorized.

Table 2. Candidate Tags.

mountains	alps	italy	mood	moody	fantasy
sunset	wilderness	adventure	peaks	bled	sunrise
lake	fog	austria	mountain	castle	river
photographer	frozen	ridge	scenery	landscape	scenic
tree	dawn	heritage	clouds	travel	reflection
water	sky	storm	summer	outdoors	nature

5.2 Case study for AP clustering

We select the target user id is 29507649@N02 for a case to analyze. After preprocessing the target user's tags of and we get the candidate tags set is shown in Table 2. In this work, we use the Affinity Propagation (AP) clustering method to cluster the candidate tags and get three clusters are shown in Table 3. We calculate the distance between the cluster center and other tags in the same cluster by the Euclidean distance (Fig. 3). We use metrics to evaluate AP clustering, and the clustering effect is shown in the Table 4.

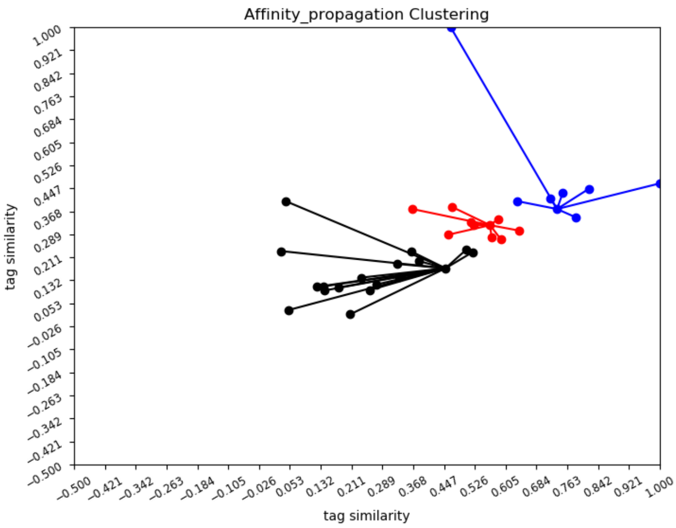

Fig. 3. Candidate tags for AP clustering.

Table 3. Tags distribution after clustering

Cluster center	Candidate tags
scenic	fog, castle, scenery, landscape, tree, heritage, clouds, sky, storm
lake	river, mountains, ridge, wilderness, peaks, alps
sunset	italy, mood, moody, fantasy, adventure, bled, sunrise, austria, photographer, frozen, dawn, water, travel, reflection, summer, outdoors, nature

Table 4. Metrics

Homogeneity	Completeness	V-measure	Mutual Information	Silhouette Coefficient
0.902	0.346	0.500	0.338	0.490

5.3 Evaluation Methodology

We use several reference methods to show the advantage of our proposed method in friend recommendation. These are: (1) Pure tag similarity. (2) Online collaborative filtering(OLCF). (3) Relational Domain Recommendation(RDR) [14].

The first is the simple tag similarity comparison. We recommend the friends of each user purely on the tag and image feature similarity. The second method is OLCF. The collaborative filtering method is widely used in recommender systems. It fills the blank entries of the user-item matrix.

Experimental results are shown in Table 5. The performance measured in terms of precision and recall. The number of potential friends recommended to a user is 10, 20 and 50 respectively. We can see Similarity have poor performance, and FRFP is better than the other three methods.

Table 5. Precision and Recall on friend recommendation

Method	Precision@10	Precision@20	Precision@50	Recall@10	Recall@20	Recall@50
Similarity	0.1037	0.0973	0.0431	0.0612	0.0732	0.1045
ROR	0.1689	0.1263	0.04871	0.0701	0.0894	0.1438
OLCF	0.2034	0.1747	0.0814	0.0839	0.1274	0.1843
FRFP	**0.2714**	**0.2347**	**0.1749**	**0.1042**	**0.1464**	**0.2147**

Precision is defined as follows:

$$Precision = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|} \qquad (13)$$

Recall is defined as follows:

$$Recall = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|} \qquad (14)$$

where U is a set of all users. $R(u)$ is the list of potential friends whom recommended to user u by the algorithm. $T(u)$ is the list of friends in test data.

To further validate the performance of FRFP, the number of potential friends is validated in more details. N is the number of potential friends recommended to target user. In our experiments, the value of N$\in$ [5, 50]. The results in terms of precision in Top-N are shown in Fig. 4, we can see that FRFP is still better than other three methods in precision when N$\in$ [5, 50]. The results in terms of recall in Top-N are also shown in Fig. 5, FRFP is observably better than other three methods in recall when N $\in$ [5, 50].

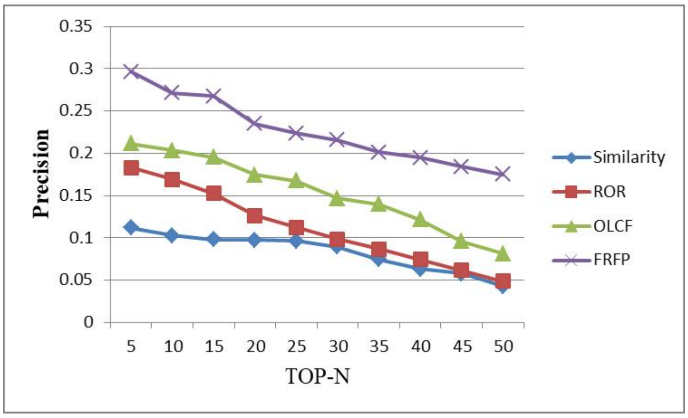

Fig. 4. Comparison of methods in precision

6 Summary and Future Work

In this paper, we propose a personalized friend recommendation method called FRFP based on user fine-grained preferences feature tag in photography community. In tagging system, we use the value of R as a score in the user-item rating

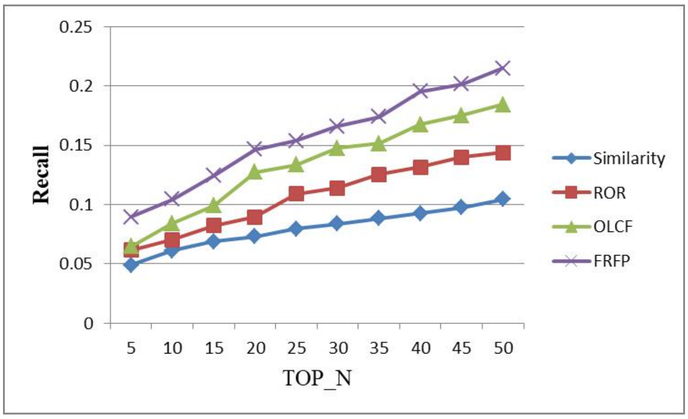

Fig. 5. Comparison of methods in recall

matrix to caculate the similarity of users' fine-grained preference. We also used Au to weight the friend recommendation list, and improving the quality of friend recommendations. However, due to the user's cold start problem in the photography community, and the user's photography preferences will change over time. For example, user uploaded new photos. Therefore, we need to deal with the above problems in the future work, and further study the hybrid recommendation system.

References

1. http://mattmahoney.net/dc/enwik9.zip.
2. https://secure.flickr.com/.
3. Frey *, Brendan J, and Dueck Delbert. Clustering by passing messages between data points. *Science*, 315(5814):972–976, 2007.
4. Mislove A, Marcon M, Gummadi K, Druschel P, and Bhattacharjee B. Measurement and analysis of online social networks. *Association for Computing Machinery (ACM)*, 2007.
5. Sarwar Badrul, Karypis George, Konstan Joseph, and Riedl John. Item-based collaborative filtering recommendation algorithms. *Proceedings of ACM World Wide Web Conference*, 2001.
6. Hwang C.-L and Yoon K.Paul. *Multiple attribute decision making. Methods and applications. A state-of- the-art survey.* 1981.
7. Neil Zhenqiang Gong, Ameet Talwalkar, Lester Mackey, Ling Huang, Eui Chul Richard Shin, Emil Stefanov, Elaine, Shi, and Dawn Song. Jointly predicting links and inferring attributes using a social-attribute network (san). *Computer Science*, 2011.
8. Mark S. Granovetter. The strength of weak ties. *American Journal of Sociology*, pages 1360–1380, 1973.
9. Siyao Han and Yan Xu. Friend recommendation of microblog in classification framework: Using multiple social behavior features. In *International Conference on Behavior*, 2015.

10. J. Hannon, M. Bennett, and B. Smyth. Recommending twitter users to follow using content and collaborative filtering approaches. *Recommender Systems*, pages 199–206, 2010.
11. Mohammed Mehedi Hasan, Noor Hussain Shaon, Ahmed Al Marouf, Md. Kamrul Hasan, and Md. Mohiuddin Khan. Friend recommendation framework for social networking sites using users online behavior. In *International Conference on Computer and Information Technology*, 2015.
12. Chaobo He, Hanchao Li, Fei Xiang, Tang Yong, and Zhu Jia. A topic community-based method for friend recommendation in online social networks via joint nonnegative matrix factorization. In *Third International Conference on Advanced Cloud and Big Data*, 2016.
13. Shangrong Huang, Jian Zhang, Shiyang Lu, and Xian Sheng Hua. Social friend recommendation based on network correlation and feature co-clustering. 2015.
14. M. Jiang, P. Cui, F. Wang, Q. Yang, W. Zhu, and S. Yang. Social recommendation across multiple relational domains. *Conference on Information and Knowledge Management*, 2012.
15. Quoc V. Le and Tomas Mikolov. Distributed representations of sentences and documents. In *International Conference on International Conference on Machine Learning*, 2014.
16. Zhepeng Li, Xiao Fang, and Olivia Sheng. A survey of link recommendation for social networks: Methods, theoretical foundations, and future research directions. *Social Science Electronic Publishing*, 2015.
17. Aditya Krishna Menon and Charles Elkan. Link prediction via matrix factorization. In *European Conference on Machine Learning and Knowledge Discovery in Databases*, 2011.
18. Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of sentences and documents. 2016.
19. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. 2013.
20. Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. 1999.
21. Shuang Qiu, Jian Cheng, Ting Yuan, Cong Leng, and Hanqing Lu. Item group based pairwise preference learning for personalized ranking. pages 1219–1222, 2014.
22. Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Conference on Uncertainty in Artificial Intelligence*, 2009.
23. BuXiao Wu, Jing Xiao, and JieMin Chen. Friend recommendation by user similarity graph based on interest in social tagging systems. 2015.
24. Xie Xing. Potential friend recommendation in online social network. In *IEEE/ACM Intl Conference on Green Computing and Communications and Intl Conference on Cyber*, 2010.
25. Huanyang Zheng and Wu Jie. Friend recommendation in online social networks: Perspective of social influence maximization. In *2017 26th International Conference on Computer Communication and Networks (ICCCN)*, 2017.