



A Cognitive Approach to Localized Image Search from the multimedia websites

D.Venkata Balakrishna¹, J.Vamsi Nath²

Department of CSE , PBRVITS Kavali¹

Assoc. Prof Department of CSE, PBRVITS Kavali²

Abstract: This generation has potential usage of social networking web sites, like face book and picasa, filkcr, YouTube allow users to tag photos, add comment, share the data. The large-scale net genies meta-data not only facilitate users in using and arranging multimedia content, but provide useful statics to improve content searching and management. Personalized search serves as one of such examples where the web search experience is improved by generating the returned list according to the modified user search intents.

In this paper, we exploit the social annotations and propose a novel framework simultaneously considering the user and query relevance to learn to personalized image search. The basic premise is to embed the user preference and query-related search intent into user-specific topic spaces. Since the users' original annotation is too sparse for topic modeling, we need to enrich users' annotation pool before user-specific topic spaces construction. The proposed framework contains two components: 1) A Ranking based Multi-correlation Tensor Factorization model is proposed to perform annotation prediction, which is considered as users' potential annotations for the images; 2) We introduce User-specific Topic Modeling to map the query relevance and user preference into the same user-specific topic space. For performance evaluation, two resources involved with users' social activities are employed. Experiments on a large-scale Flickr dataset demonstrate the effectiveness of the proposed method.

Keywords: Personalized image search, tensor factorization, topic model, social annotation.

I. INTRODUCTION

Keyword-based search has been the most popular search paradigm in today's search market. Despite simplicity and efficiency, the performance of keyword-based search is far from satisfying. Investigation has indicated its poor user experience - on Google search, for 52% of 20,000 queries, searchers did not find any relevant results [1]. "IR" has the interpretation of both information retrieval and infra-red. 2) Users may have different intentions for the same query, e.g., searching for "jaguar" by a car fan has a completely different meaning from searching by an animal specialist. One solution to address these problems is personalized search, where user-specific information is considered to distinguish the exact intentions of the user queries and re-rank the list results. Given the large and growing importance of search engines, personalized search has the potential to significantly improve searching experience. Compared with non-personalized search, in personalized search, the rank of a document (web page, image, video, etc.) in the result list is decided not only by the query, but by the preference of user. Fig. 1 shows a toy example for non-personalized and personalized image search results. The non-personalized search returned results only based on the query relevance and displays jaguar car images as well as wild cat on the top. While personalized search consider both query relevance

and user preference, therefore the personalized results from an animal lover rank the leopard images on the top. This provides a natural two-step solution scheme.

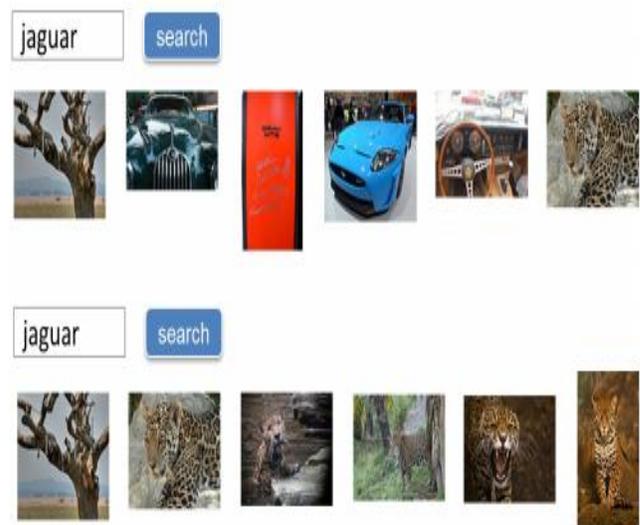


Fig. 1. Toy example for non-personalized (top) and personalized (bottom) search results for the query "jaguar".



Most of the existing work [2], [3], [4], [5] follow this scheme and decompose personalized search into two steps: computing the non-personalized relevance score between the query and the document, and computing the personalized score by estimating the user's preference over the document. After that, a merge operation is conducted to generate a final ranked list. While this two-step scheme is extensively utilized, it suffers from two problems. 1) The interpretation is less straight and not so convinced. The intuition of personalized search is to rank the returned documents by estimating the user's preference over documents under certain queries. Instead of directly analyzing the user-query-document correlation, the existing scheme approximates it by separately computing a query-document relevance score and a user-document relevance score. 2) How to determine the merge strategy is not trivial. In this paper, we simultaneously considers the user and query dependence and present a novel framework to tackle the personalized image search problem To investigate on user preference and perform user modeling, the popular social activity of tagging is considered.

Collaborative tagging has become an increasingly popular means for sharing and organizing resources, leading to a huge amount of user-generated annotations. Online photo sharing websites, such as Flickr, Picasa, Zoomr and Interest allow users as owners, taggers, or commenters for their contributed contents to interact and collaborate with each other in a social media dialogue. Various researchers have investigated the applicability of social annotations to improve web search [6], [7], [8], [9], [10]. Recently, social annotations are employed for automatic evaluation of personalized search [2], [11], [3]. A fundamental assumption is that, *the users' tagging actions reflect their personal relevance judgment*. For example, if a user tagged "festival" to an image, it is probable that the user will consider this image as relevant if he/she issues "festival" as a query. Illustrated by this, the intuition of this paper is that if the users' annotations to the images are available, we can directly estimate the users' preference under certain queries. The fact is that the original annotation available is not enough for user preference mining. Therefore, we transfer the problem of personalized image search to users' annotation prediction. Moreover, as queries and tags do not follow simple one-to-one relationship, we build user-specific topic spaces to exploit the relations between queries and tags.

A. Framework

The framework of this paper is shown in Fig.2. It contains two stages: offline model training stage and online personalized search response stage. Typically a weighting parameter will be optimized to balance the two scores [2], or the learnt user preference is used to re-rank the query

relevance-based original list [5]. For the offline stage, three types of data including users, images and tags as well as their ternary interrelations and intra-relations are first collected. We then perform users' annotation prediction. Many methods [12], [13], [14] for tag recommendation and prediction have been proposed in social bookmark sites, e.g., Bibsonomy, Del.icio.us, Last.fm, etc.

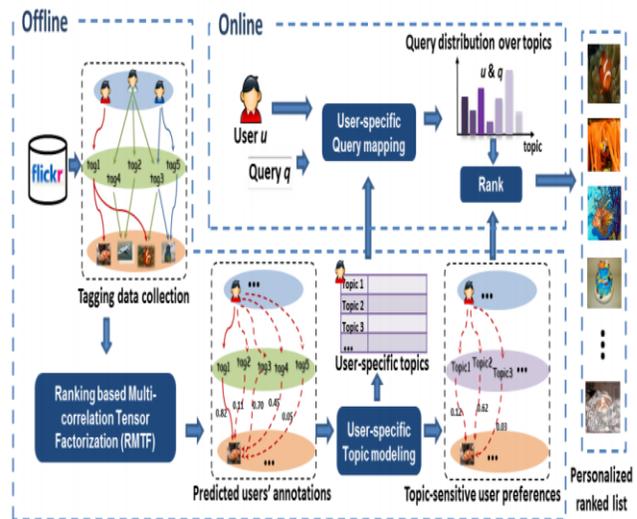


Fig. 2. The proposed framework.

Since the photo sharing websites utilize a different tagging mechanism that repetitive tags are not allowed for unique images, besides the common noisy problem, it has more severe sparsity problem than other social tagging systems. To alleviate the sparsity and noisy problem, we present a novel method named Ranking based Multi-correlation Tensor Factorization (RMTF) to better leverage the observed tagging data for users' annotation prediction. Zhu *et al.* [15] has demonstrated that the semantic space spanned by image tags can be approximated by a smaller subset of salient words from the original space.

Illustrated by this, we employ lowrank approximation to extract the compact representation for image, tag and user, and at the same time reconstruct the user-image-tag ternary relations for annotation prediction. With the observed user-tag-image ternary relations as input, the reconstructed ternary relations can be viewed as users' potential annotations for the images. Following the assumption we mentioned in the introduction, we can straightly utilize the predicted user annotations for personalized image search, i.e., if a user has a high probability to assign the tag t to an image, the image should be ranked higher when the user issues query t . However, this formulation has two problems. It is unreasonable to assign the query to a single tag in the tag vocabulary, e.g., when a user searches "cheerdance", he/she would like the images that 2 We use Flickr, the



popular photo sharing website, for our experiments. The information of users, images and tags can be acquired from the Flickr API: <http://www.flickr.com/services/api>. We show a running example consisting of three users, five tags and four images. We conduct the sparsity investigation between Flickr, the publication tagging website Bibsonomy and the music sharing websites Last.fm in Section V.

He/she annotated with semantic related tag “cheerleader” are also ranked higher. 2) There are variations in individual user’s tagging patterns and vocabularies, e.g., the tag “jaguar” from an animal specialist should be related to “leopard”, while a car fan will consider “jaguar” more related to “autos”. To address the two problems, we perform *User-specific Topic Modeling* to build the semantic topics for each user. The user’s annotation for an image is viewed as *document*. The individual tag to the image is *word*. User’s annotations for all the images constitute the *corpus*. As the original annotation is too sparse for topic modeling, we use the reconstructed ternary relations as the document collections. The user’s topic distribution per image can be considered as his/her preference over the image on the learned user-specific topic space. Therefore, after the offline stage, two outcomes are stored in the system, the user-specific topics and topic-sensitive user preferences. For the online stage, when a user u submits a query q , we first map the query q to user u -specific topics. The query distribution is then sent to the rank module and employed as the weight on topics to calculate the user u ’s topic sensitive preferences over the images. Finally, the images are ranked according to the calculated user’s preferences, which simultaneously consider the query and user information. The contributions of this paper are summarized as three folds.

- We propose a novel personalized image search framework by simultaneously considering user and query information. The user’s preferences over images under certain query are estimated by how probable he/she assigns the query-related tags to the images.
- A ranking based tensor factorization model named RMTF is proposed to predict users’ annotations to the images.
- To better represent the query-tag relationship, we build user-specific topics and map the queries as well as the users’ preferences onto the learned topic spaces.

2. RELATED WORK

In recent years, extensive efforts have been focusing on personalized search. Regarding the resources they leveraged, explicit user profile [17], relevance feedback [18], user history data (browsing log [19], click-through data [20], [21] and social annotations [11], [8], [4] etc.), context information [23] (time, location, etc.) and social network [1], [3], [16] are exploited. For the implementation there are two

primary strategies [24], query refinement and result processing. In the following we review the related work by the strategy they used.

Query Refinement

also called Query Expansion, refers to the modification to the original query according to the user information. It includes augmenting the query by other terms [18], [25] and changing the original weight of each query term [26]. Kraft *et al.* [18] utilized the search context information collected from users’ explicit feedback to enrich the query terms. Chirita *et al.* [25] proposed five generic techniques for providing expansion terms, ranging from term and expression level analysis up to global co-occurrence statistics and external thesauri. While, Teevan *et al.* [26] re-assigned the weights of original query terms using BM25 weighting scheme to incorporate user interests as collected by their desktop indexes. We do not explicitly perform query refinement in this paper. However, mapping the queries into user-specific topic spaces can be considered as implicit query refinement.

Result Processing

can be further classified into result filtering and re-ranking. Result filtering aims to filter irrelevant results that are not of interest to a particular user [27]. While, result re-ranking focuses on re-ordering the results by the degree of users’ preferences estimated. Since our work falls into this category, we mainly review the related work on result re-ranking. Chirita *et al.* [17] conducted an early work by reranking the search results according to the cosine distance between each URL and user interest profiles constructed. Qiu *et al.* [21] extended Topic-Sensitive PageRank by incorporating users’ preference vectors. By aggregating the search results from multiple search engines, Liu *et al.* [22] introduced a new method for visual search reranking called CrowdReranking.

A typical work is performed by Xu *et al.* [2], in which the overall ranking score is not only based on term similarity matching between the query and the documents but also topic similarity matching between the user’s interests and the documents’ topics. In the similar spirits, Cai [4] formalized query and user relevance measurement separately as fuzzy requirement satisfaction problem. Lu *et al.* [5] utilized a co-clustering method to extract latent interest dimensions, and re-rank the images by combining latent interest based user preference and query relevance.

In our work, there is also a topic space to model user preference. However, regarding the variations in user’s tagging vocabularies, we build user-specific topics and calculate topic-sensitive user preference over images, which



differentiate our work from Xu [2] and Lu's [5]. Besides, most of the existing work decompose the overall ranking score into query relevance and user preference and generate two separated ranked list. While in this paper, we map the queries into the same user-specific topic space and directly compute the users' preference under certain queries.

3.RANKING BASED MULTI-CORRELATION TENSOR FACTORIZATION

In this section, we present the algorithm for annotation prediction. Table I lists the key notations used in this paper. There are three types of entities in the photo sharing websites. The tagging data can be viewed as a set of triplets. Let denote the sets of users, images, tags and the set of observed tagging data is denoted i.e., each triplet $(u; i; t)$ means that user u has annotated image i with tag t . The ternary interrelations can then constitute a three dimensional tensor Fig.6(a) shows the tensor constructed from the running example in Fig.2. Predicting the users' annotations to the images are related to reconstructing the user-tag-image ternary interrelations. We use Tucker decomposition [31], a general tensor factorization model, to perform the low-rank approximation. In Tucker decomposition, the tagging data Y are estimated by three low rank matrices and one core tensor: where \times_n is the tensor product of multiplying a matrix on mode n . Each matrix corresponds to one factor. The core tensor contains the interactions between the different factors.

The ranks of decomposed factors are denoted by and this is called *rank*Tucker decomposition. Under Tucker decomposition, we need to design appropriate objective function to optimize the latent factors $U; I; T; C$ and then calculate the reconstructed tensor by Eq.2. In this paper, a model named RMTF is proposed to design the objective function. To better leverage the observed tagging data, we first introduce a novel ranking based optimization scheme for representation of the tagging data. Then the multiple intra-relations among users, images and tags are utilized as the smoothness constraints to tackle the sparsity problem.

A. Ranking based Optimization Scheme

A direct way to approximate Y is to minimize the sum of point-wise loss on \hat{Y} : \min As this optimization scheme tries to fit to the numerical values of 1 and 0, we refer it as the *0/1 scheme*. However, under the situation of social image tagging data, the semantics of encoding all the unobserved data as 0 are incorrect, which is illustrated with the running example:

Firstly, the fact that *user3* has not given any tag to *image2* and *image4* does not mean *user3* considering all the tags are

bad for describing the images. Maybe he/she does not want to tag the image or has no chance to see the image.

Secondly, *user1* annotates *image1* with only tag3. It is also unreasonable to assume that other tags should not be annotated to the image, as some concepts may be missing in the user-generated tags and individual user may not be familiar to all the relevant tags in the large tag vocabulary. According to the optimization function in Eq. *scheme* tries to predict 0 for both cases. To address the above two 5 we call the triplets like as neutral triplets. Issues, in this paper, we present a ranking optimization scheme which intuitively takes the user tagging behaviors into consideration. Firstly, we note that only the qualitative difference is important and fitting to the numerical values of 1 and 0 is unnecessary.

Therefore, instead of solving an point-wise classification task, we formulate it as a ranking problem which uses tag pairs within each user-image combination as the training data and optimizes for correct ranking. We provide some notations for easy explanation. Each user image combination $(u; i)$ is defined as a *post*. The set of observed posts is denoted as (4) Note that the ranking optimization is performed over each post and within each post a positive tag set and a negative tag set are desired to construct the training pairs. We assume that any tag is a better description for image It is arbitrary to treat the neutral triplets as either positive or negative and we remove all the triplets in M from the learning process (filled by bold question marks in .

We then consider two characteristics of the user tagging behaviors to choose and On one hand, some concepts may be missing in the user-generated tags. We assume that the tags co-occurring frequently are likely to appear in the same image (we call it context-relevant). On the other hand, users will not bother to use all the relevant tags to describe the image. The tags semantic-relevant with the observed tags are also the potential good descriptions for the image.

On the other hand, users will not bother to use all the relevant tags to describe the image. The tags semantic-relevant with the observed tags are also the potential good descriptions for the image. The tags with the k -highest affinity values are considered semantic-relevant and context-relevant To perform the idea, we build a tag affinity graph based on tag semantic and context intra-relations (detailed in Section III.B).The affinity graphs can be utilized as the regularization terms to impose smoothness constraints for the latent factors.

Take the image affinity graph WI and the image factor matrix I as example, the regularization term is. We can build similar regularization terms for the user and tag factors. Combining with Eq.6, we obtain the following overall objective function.In this phase the relevant images of user



query are ranked on the basis of the popularity of the image. The user generated metadata through their everyday activity on photo sharing websites is useful to get the popularity of the image on basis of which the ranking model works. The result will give the priority to that image which is most popular among the search result.

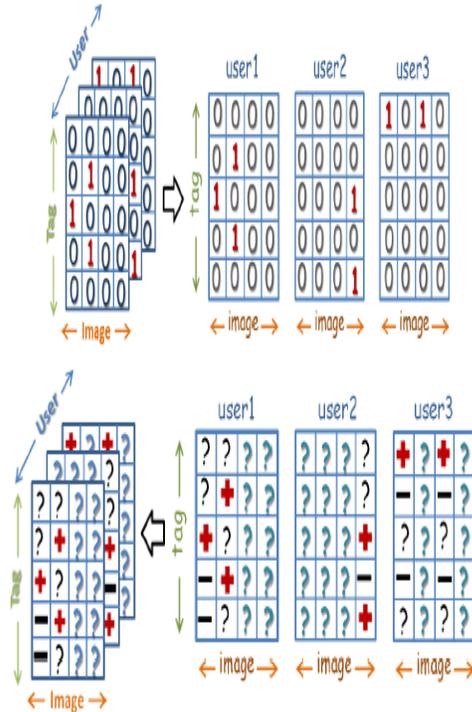


Fig. 3. Tagging data interpretation. (a) 0/1 scheme (b) ranking scheme

B. Multi-correlation Smoothness Constraints

Photo sharing websites differentiate from other social tagging systems by its characteristic of self-tagging: most images are only tagged by their owners. Fig.4 (a) shows the #tagger statistics for Flickr and the webpage tagging system Del.icio.us. We can see that in Flickr, 90% images have no more than 4 taggers and the average number of tagger for each image is about 1.9. However, the average tagger for each webpage in Del.icio.us is 6.1. The severe sparsity problem calls for external resources to enable information propagation. We assume that two items with high affinities should be mapped close to each other in the learnt factor subspaces. In the following, we first introduce how to construct the tag affinity graph, and then incorporate them into the tensor factorization framework. To serve the ranking based optimization scheme, we build the tag affinity graph based on the tag semantic relevance and context relevance.

The context relevance of tag *and* is simply encoded by their weighted co-occurrence in the image collection. The user affinity graph and image affinity graph are constructed based on statistics of co-joined groups and visual similarity,

respectively. However in practice, the queries and tags do not follow oneto-one relationship - one query usually corresponds to several related tags in the tag vocabulary. Besides, the query-tag correspondence differs from user to user. Therefore, we build topic spaces for each user to exploit this user-specific one-to-many relationship. We investigate on a Flickr dataset of 270K images that the average number of annotated images per user is only 30. The detailed distribution is shown in .The individual tag is viewed as *word*, while the user’s annotation to one image corresponds to one document.

Table-1
 THE DOMINATING USER-SPECIFIC TOPICS FOR TWO EXAMPLE USERS.

| | | |
|---------------|----------------|--|
| User A | Topic 1 | military, aircraft, battleship, navy, artillery, iraq, aircraftcarrier, barracks |
| | Topic 2 | apple, computer, art, girl, cellphone, cool, vintage, digital |
| | Topic 3 | athlete, basketball, baseball, actor, sports, art, film, black |
| User B | Topic 1 | buddha, buddhist, temple, religion, buddhism, thailand, asia, ancient |
| | Topic 2 | blossoms, blooms, nature, macro, flower, bravo, butterfly, spring |
| | Topic 3 | airplane, boeing, aircraft, airport, aviation, jet, aeroplane, cockpit |

When user *A* searches “aircraft”, the images likely to be annotated by military-related tags are ranked higher according to Eq.18. While, when user *B* searches “aircraft”, the images likely to be annotated by aerocraft-related tags will be ranked higher. We can see that the query relevance and user preference are simultaneously integrated into this personalized formulation.

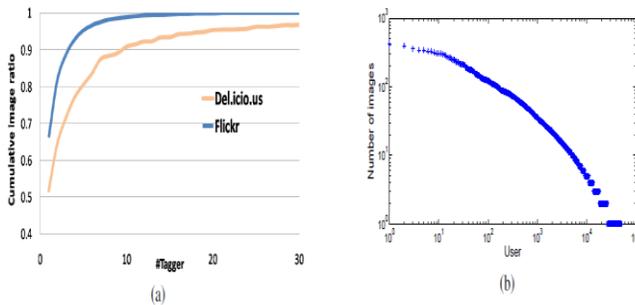
A. Dataset

We perform the experiments on a large-scale web image dataset, It contains 269,648 images with 5,018 unique tags collected from Flickr. We crawled the images’ owner information and obtained owner user ID of images. We only consider the one word-based queries in this paper and handling complex queries of multiple words is our future work.10 Due to link failures, the owner ID of some images is unavailable.

TABLE II
 DATASET STATISTICS FOR SEVERAL TAGGING SYSTEMS



| Dataset | User jU_j | Tag jT_j | Document jD_j | Observed Tagging Triplet jO_j |
|----------------|-------------|------------|-----------------|---------------------------------|
| Bibsonomy [13] | 116 | 412 | 361 | 10,148 |
| Last.fm [34] | 2,917 | 1,853 | 2,045 | 219,702 |
| NUS-WIDE | 50,120 | 5,018 | 247,849 | 4,541,326 |
| NUS-WIDE15 | 3,372 | 5,018 | 124,099 | 1,223,254 |

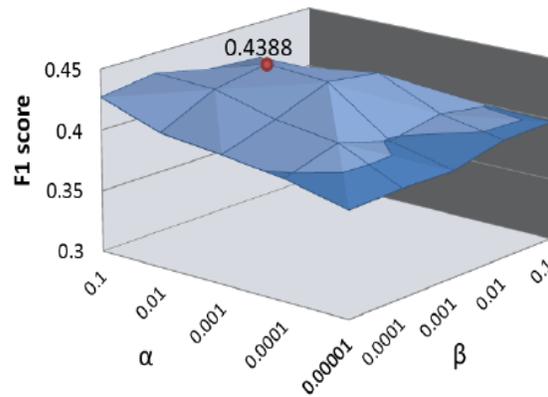


B. Parameter Setting

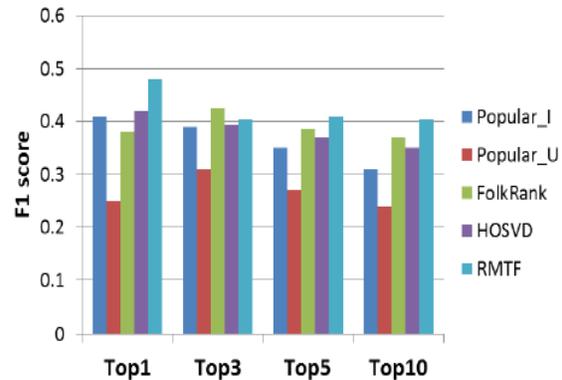
NUS-WIDE15 is randomly split into two parts, 90% for training and testing (denoted as S), and 10% for validation (denoted as V). The result of annotation prediction directly affect the performance of personalized search. In our work, we select parameters according to the performance of annotation prediction. The evaluation for annotation prediction is detailed in the next subsection. relatively steady when α and β change within a certain range. We set $\alpha = 0:01$ and $\beta = 0:001$, which achieves the highest average F1 score. The most important parameter for userspecific topic modeling is the number of latent topics for each user. For now the number is set same for different users and $K = 20$. We investigate the influence of K in the following experiment.

C. Annotation Prediction

We propose the novel RMTF model for users' annotation prediction. In this subsection, we first evaluate the performance of RMTF for annotation prediction. Following the evaluation process from [13], for each user we randomly remove all triplets he/she has annotated for one image to constitute the test set $Stest$ – i.e., we remove one post for each user. The remaining observed user-image-tag triplets are used for regularized tensor factorization. Then we learn the model and predict top-N lists for each of the removed posts PS_{test} based on the reconstructed tensor. We compute the recall and precision of the top-N recommended tags and report the F1 score of the average recall and precision.



Impact of parameters α and β .



The F1 score of annotation prediction for different methods.

Four annotation prediction methods are performed for comparisons: most popular tags for image (Popular I), most popular tags for user (Popular U), FolkRank [13] and HOSVD [14]. Fig. 6 illustrates the results. It is shown that RMTF generally performs the best, and with the increasing number of recommended tags, the F1 score decreases less steeper for RMTF than the other methods. This coincides with our discussions in the introduction that the proposed ranking scheme as well as exploiting the tag semantic-and-context relevance better alleviates the severe sparsity and noisy problem for Flickr dataset.

D. Personalized Search

In the research community of personalized search, evaluation is not an easy task since relevance judgement can only be evaluated by the searchers themselves. The most widely accepted approach is user study [24], [26], [17], [25], where participants are asked to judge the search results.



Obviously this approach is very costly. In addition, a common problem for user study is that the results are likely to be biased as the participants know that they are being tested. Another extensively used approach is by user query logs or click through history [20], [35]. However, this needs a large-scale real search logs, which is not available for most of the researchers.

Social sharing websites provide rich resources that can be exploited for personalized search evaluation. User's social activities, such as rating, tagging and commenting, indicate the user's interest and preference in a specific document. Recently, two types of such user feedback are utilized for personalized search evaluation. The first approach is to use social annotations [2], [11], [3]. The main assumption behind is that the documents tagged by user u with tag t will be considered relevant for the personalized query ($u; t$). Another evaluation approach is proposed for personalized image search on Flickr [5], where the images marked *Favorite* by the user u are treated as relevant when u issues queries. The two valuation approaches have their pros and cons and supplement for each other. We use both in our experiments and list the results in the following.

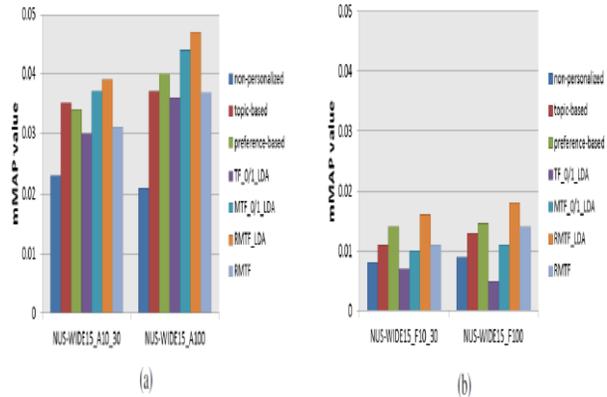
1. Annotation-based Evaluation

We follow Xu's [2] evaluation framework and first compare the performances according to users' original annotations. To perform the evaluation in the situations of users with different amount of original annotations, we build two test scenarios: 1) 30 randomly selected users who tagged 10-30 images and their tagging records, denoted as NUS-WIDE15 A10 30. 2) 30 randomly selected users who tagged more than 100 images and their tagging records, denoted as NUS-WIDE15 A100. For NUSWIDE15 A100, the overlapping 18 tags the 50 users used are selected as the test queries, while for NUS-WIDE15 A10 30, the number of test queries is 11. The statistics of the testing sets are shown in Table IV. In order to reduce the dependency between original annotations and evaluation, we remove the tagging data related to the test queries. It is done as follows: for each personal query we remove the triplets from the training set.

TABLE-III
 TESTING SET STATISTICS FOR EVALUATION

| Testing set | # User | # Query | # Images tagged/favorited | # Tags annotated |
|-------------------|--------|---------|---------------------------|------------------|
| NUS-WIDE15 A10 30 | 30 | 11 | 253 | 14,148 |
| NUS-WIDE15 A100 | 30 | 18 | 4566 | 319,702 |
| NUS- | 30 | 15 | 233 | 5,015 |

| WIDE15 F10 30 | | | | |
|-----------------|----|----|-------|--------|
| NUS-WIDE15 F100 | 19 | 15 | 31,24 | 19,254 |



The mMAP value of personalized search for different methods (a) Annotation-based; (b) Favorite-based

The results are shown in Fig.7(a). *Non-personalized* denotes the non-personalized rank result by only considering the query relevance. We can see that all the personalized methods outperform the non-personalized scheme. Comparing between the two test scenarios of NUS-WIDE15 A10 30 and NUSWIDE15 A100, the performances of personalized methods improve as the test users' original annotations increase. This is reasonable as these methods utilize the social annotation resources and the more user feedback is available, the more accurate user preferences can be estimated. What is interesting is that the preference-based model [5] and the proposed model are more sensitive to the amount of original annotations. The reason may be that [5] and our methods extract topic spaces by explicitly exploiting the tagging data, while in the topicbased model [2], the topic space is pre-defined and the original annotation is just used to generate the topic vector.

2) Favorite-based Evaluation

There is a delicate issue with annotation-based evaluation. Both the input to the personalized models and the evaluation for the output results are based on the original annotations. Although the specific tagging data have been removed when testing the personal query), as individual user's tagging vocabulary ends to be limited, the remaining annotations will implicitly provide the association between u and t . For example, assuming one user u usually tag "wildlife" and "animal" together, when he/she issues "wildlife" as test query, though all have been removed from the training process, regarding "wildlife" and "animal" are likely to have a close relation in the userspecific topics, the images tagged by "animal" will be given high probability



and guide the final rank. On Flickr, users are encouraged to express their preference on images by adding *Favorite* marks. Illustrated by Lu's evaluation framework [5], we employ users' *Favorite* marks for evaluation, which are not used in the training process. This guarantees that personalization is evaluated without any prior knowledge.

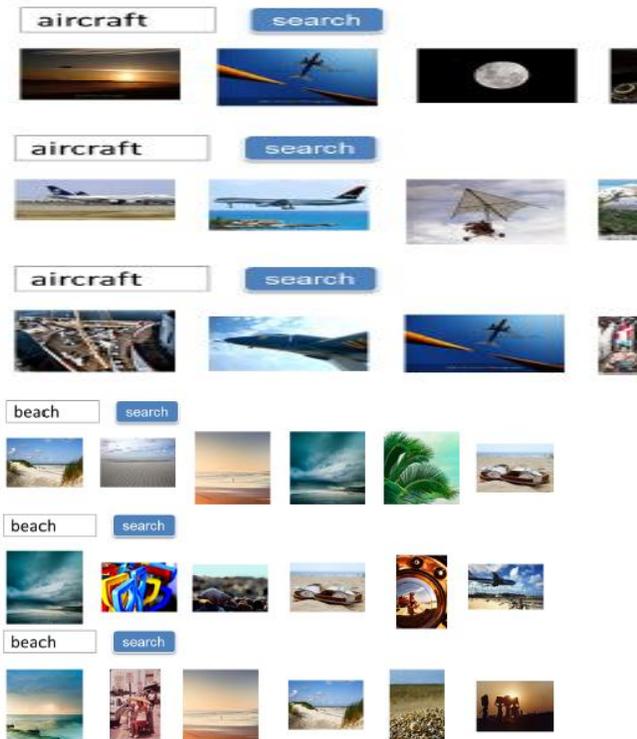
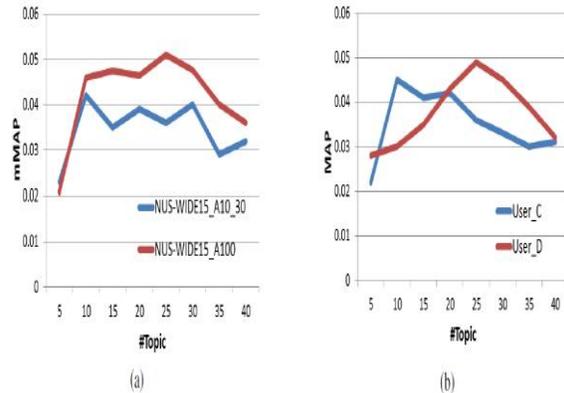


Fig. 4. Example of non-personalized (top) and personalized (middle for User A and bottom for User B) search results for query

Displays exemplary search results for the query "aircraft". The top six non-personalized results and the personalized results of User A and User B mentioned in Section IV.A are shown. We can see that by simultaneously considering query relevance and user information, the proposed RMTF LDA captures the user's preference under certain topics. As a result of mapping "aircraft" to Topic 1 of Table II, the top search results for user A mainly focus on aerocrafts. While, for user B, the top search results are basically military related, which coincides with user B's preference. For the baseline method which separate query relevance and user preference, sometimes the search results are hard to interpret.

For example, the second and third images for user B in Fig. 8(a) are ranked higher because user B has a major interest in religion and flower. However, these images have little relation with aircraft. We note that for some general queries which have clear search intents, personalized search tends to fail. Fig. 9 illustrates one of such examples. With "beach"

having common understanding to different users, incorporating user information will generate confusing search results. There are literatures [35], [36] discussing the issue about when to perform personalization. It seems that the benefit of personalization is highly dependent on the ambiguity of the query.



The influence of topic number K (a) mMAP for all test users; (b) MAP for two users from NUS-WIDE15 A100

Fig. 10. The influence of topic number K (a) mMAP for all test users; (b) MAP for two users from NUS-WIDE15 A100. Since there is no conclusion to this problem, in this paper we focus on the problem of how to perform personalization and discussion of when to perform personalization is beyond the scope of this project.

3) Influence of Topic Number K

For the standard LDA, the number of latent topics needs to be specified. In the above experiments, we set the same number of topics for all users and $K = 20$. In the following, we variate the selection of K and investigate the influence of topic number. We utilize the annotation-based evaluation scheme to compute mMAP. The results are illustrated in Fig.10. It is shown that mMAP displays no definite trend as K changes (Fig.10(a)), while for individual user C and user D there exist obvious optimal K (See Fig.10(b), obviously user C and D have an optimal $K = 10$ and $K = 25$ respectively). This observation is inline with the expectations that users have different topic spaces and validate the necessity of user-specific topic modeling. In addition, user-specific topic number should be specified in the future work. There is a number of extension work on standard LDA to automatically select the number of topics. The most common one is HDP-LDA, which uses Hierarchical Dirichlet Processes (HDP, [37]) to model the Dirichlet mixtures in LDA.

4. CONCLUSION

How to effectively utilize the rich user metadata in the social sharing websites for personalized search is challenging as well as significant. In this paper we propose a novel framework to exploit the users' social activities for personalized image search, such as annotations and the participation of interest groups. The query relevance and user preference are simultaneously integrated into the final



rank list. Experiments on a large-scale Flickr dataset show that the proposed framework greatly outperforms the baseline.interest profile. Therefore, this framework can be extended to any applications based on interest profiles. 3) For batch of new data (new users or new images), we directly restart the RMTF and user-specific topic modeling process. While, for a small amount of new data, designing the appropriate update rule is another future direction. 4) Utilizing large tensors brings challenges to the computation cost. We plan to turn to parallelization (e.g. parallel MATLAB) to speedup the RMTF converge process. Moreover, the distributed storing mechanism of parallelization will provide a convenient way to store very large matrices and further reduce the storage cost.

REFERENCES

- [1] B. Smyth, "A community-based approach to personalizing web search," *Computer*, vol. 40, no. 8, pp. 42–50, 2007. [2] S. Xu, S. Bao, B. Fei, Z.
- [2] Su, and Y. Yu, "Exploring folksonomy for personalized search," in *SIGIR*, 2008, pp. 155–162.
- [3] D. Carmel, N. Zwerdling, I. Guy, S. Ofek-Koifman, N. Har'El, I. Ronen, E. Uziel, S. Yagev, and S. Chernov, "Personalized social search based on the user's social network," in *CIKM*, 2009, pp. 1227–1236.
- [4] Y. Cai and Q. Li, "Personalized search by tag-based user profile and resource profile in collaborative tagging systems," in *CIKM*, 2010, pp. 969–978.
- [5] D. Lu and Q. Li, "Personalized search on flickr based on searcher's preference prediction," in *WWW (Companion Volume)*, 2011, pp. 81–82.
- [6] P. Heymann, G. Koutrika, and H. Garcia-Molina, "Can social bookmarking improve web search?" in *WSDM*, 2008, pp. 195–206.
- [7] S. Bao, G.-R. Xue, X. Wu, Y. Yu, B. Fei, and Z. Su, "Optimizing web search using social annotations," in *WWW*, 2007, pp. 501–510.
- [8] D. Zhou, J. Bian, S. Zheng, H. Zha, and C. L. Giles, "Exploring social annotations for information retrieval," in *WWW*, 2008, pp. 715–724.
- [9] J. Tang, S. Yan, R. Hong, G. Qi and T. Chua, "Inferring Semantic Concepts from Community-Contributed Images and Noisy Tags," in *ACM Multimedia*, 2009, pp. 223–232.
- [10] J. Tang, H. Li, G. Qi and T. Chua, "Image Annotation by Graph-Based Inference With Integrated Multiple/Single Instance Representations," in *IEEE Trans. Multimedia*, 2010, vol. 12, no. 2, pp. 131–141, 2010.
- [11] M. J. Carman, M. Baillie, and F. Crestani, "Tag data and personalized information retrieval," in *SSM*, 2008, pp. 27–34.
- [12] R. Jäschke, L. B. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme, "Tag recommendations in folksonomies," in *PKDD*, 2007, pp. 506–514.
- [13] R. Jäschke, L. B. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme, "Tag recommendations in social bookmarking systems," *AI Commun.*, vol. 21, no. 4, pp. 231–247, 2008.
- [14] P. Symeonidis, A. Nanopoulos and Y. Manolopoulos, "A Unified Framework for Providing Recommendations in Social Tagging Systems Based on Ternary Semantic Analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 2, pp. 179–192, 2010.