International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307, Volume-4 Issue-2, May 2014

Personalized Image Search

Nihar Ranjan, Neha Thombare, Pallavi Deshmukh, Simantini Patil, Shailesh Jain

Abstract— Personalized Search is a feature in which when a user is logged into a account, all of his or her searches on Personal Search are recorded into Web History. Then, when a user performs a search, the search results are not only based on the relevancy of each web page to the search term, but the service also takes into account what websites the user previously visited through search results to determine which search results to determine for future searches, to provide a more personalized experience. The feature only takes effect after the user has performed several searches, so that it can be calibrated to the user's tastes. Social sharing websites like facebook, twitter, YouTube they are allowing user to comment, tag, like and unlike the shared documents or images. Rapid Increase in the search services for social websites has been developed.

Index Terms—Personalized Search, Tagging, Topic Model

I. INTRODUCTION

Rapidly developed social sharing websites allow users to create, share, annotate and comment medias. The large-scale user-generated meta-data not only facilitate users in sharing and organizing multimedia content, but provide useful information to improve media retrieval 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. Here, 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 component such as.

We introduce User-specific Topic Modeling to map the query relevance and user preference into the same user-specific topic space.

Preliminary experiments demonstrate the improvement of the proposed model compared to existing one-fit-all methods and a user-based collaborative filtering method.

We formulate and study search algorithms that consider a user's prior interactions with a wide variety of content to personalize that user's current Web search. Rather than relying on the unrealistic assumption that people will precisely specify their intent when searching, we pursue techniques that leverage implicit information about the user's interests. This information is used to re-rank Web search results within a relevance feedback framework.

Manuscript Received May, 2014

Nihar Ranjan, Computer Department, University of Pune, Pune , India.
Neha Thombare, Computer Engineering, University of Pune, Pune, India.
Pallavi Deshmukh , Computer Engineering, University of Pune, Pune, India.
Simantini Patil, Computer Engineering, University of Pune, Pune, India.
Shailesh Jain , Computer Engineering, University of Pune, Pune, India.

We explore rich models of user interests, built from both search-related information, such as previously issued queries and previously visited Web pages, and other information about the user such as documents and email the user has read and created. Our research suggests that rich representations of the user and the corpus are important for personalization, but that it is possible to approximate these representations and provide efficient client-side algorithms for personalizing search. We show that such personalization algorithms can significantly improve on current Web search. When searching photos by submitting a query, a user may receive hundreds or thousands of returned results, e.g., 118,147 photos are returned by searching with "*Great Wall*".

Obviously, users need a tool to assist them in getting access to interested photos more easily. Personalized search serves as such a tool which rearranges the returned results based on the preference of the searcher. Typically, users are interested in more than one field, and the searcher may share different interests with different friends. The variety of users' implicit interests can be mined and encoded into the latent interest dimensions. Friends may contribute differently to searcher's preference prediction according to the submitted query and the interest distribution. For example, a friend distributed consistently with the searcher on the latent dimensions related to *Travel* and *Landscape* will contribute much to a query like '*Great Wall*'.

II. PROPOSED FRAMEWORK

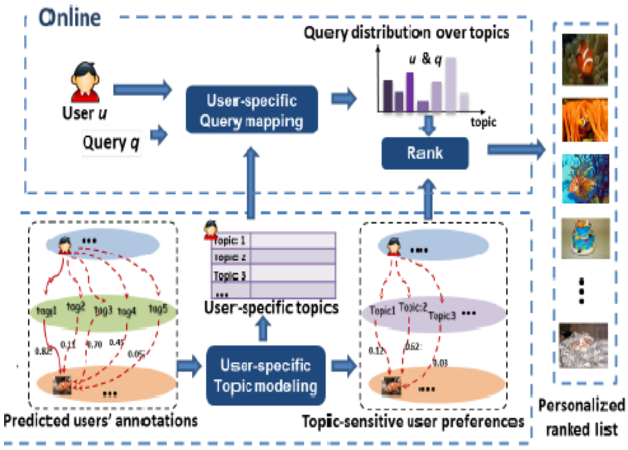
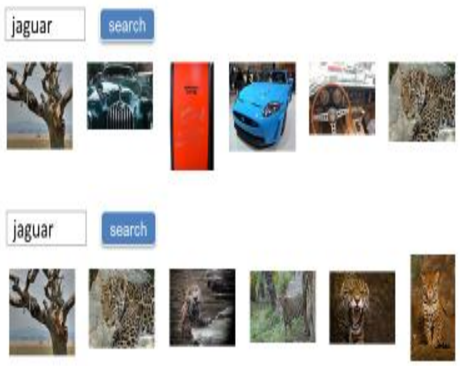


Fig .1

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. This is due to two reasons.

1) Queries are in general short and nonspecific, e.g., the query of "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.

The framework of shown in Fig.1. It contains online personalized search response stage.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.In addition to it the following problems may appear:

1) It is unreasonable to assign the query to a single tag in the tag vocabulary, e.g., when a user searches “cheer dance”, he/she would like the images that 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.. 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 topicsensitive preferences over the images. Finally, the images are ranked according to the calculated user’s preferences, which simultaneously considers the query and user information.

The contributions of this paper are summarized as threefolds:

- 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.

- 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.

III. USER-SPECIFIC TOPIC MODELING

With the reconstructed user-tag-image ternary interrelations, we can directly perform the personalized image search: when user u submits a query q, the rank of image i is inversely proportional to the probability of u annotating i with tag q: $rank(i|q; u) \propto 1$ ^$yu;i;q$ However in practice, the queries and tags do not follow one to- 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. Particularly, for each user u, the tags with 100 highest ^$yu;i;t$ are reserved as the annotations for image i. The individual tag is viewed as *word*, while the user’s annotation to one image corresponds to one *document*. We assumes that in one corpus, documents are generated from a set of K latent topics $\{topic1; \cdots ; topicK\}$. Document $\mathbf{t}i$ is the tags assigned to image i by individual user. In $\mathbf{t}i$, each word t is associated with a latent topic.

IV. BEHAVIORAL TRACKING

The main motive behind behavioral tracking is to provide the user with efficient data and to predict its annotations .The search engine includes the concept of category .Whenever the user is being logged into its account he/she has to select the category interested in eg. fruit category.This will help the user to get images of apple fruit instead of apple phone.But this work seems to manually operated.In order to make it automatic ,the concept of behavioral tracking is used.This will track the users behavior in terms of his/her likes and dislikes and his previously done searches.So whenever the user is being logged for the second time he will get the most searched images.eg if an user has mostly searched for apple fruit then according to his behavior the tracker will show him/her the fruit images.In some case if user needs apple phone images then the user will select the category of gadgets and will be shown apple phone images.

V. CONCLUSION AND FUTURE WORKS

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 dataset show that the proposed framework greatly outperforms the baseline. In the future, we will improve our current work along four directions.

1) In this paper, we only consider the simple case of one word-based queries. Actually, the construction of topic space provides a possible solution to handle the complex multiple words-based queries. We will leave it for our future work.
2) During the user-specific topic modeling process, the obtained user-specific topics represent the user’s distribution on the topic space and can be considered as user’s 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 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 distributed storing mechanism of parallelization will provide a convenient way to store very large matrices and further reduce the storage cost.

REFERENCES

1. Learn to Personalized Image Search from the Photo Sharing Websites Jitao Sang, Changsheng Xu, Senior Member, IEEE, Dongyuan Lu
2. B. Smyth, "A community-based approach to personalizing web search,"Computer, vol. 40, no. 8, pp. 42–50, 2007.
3. Personalized Search on Flickr based on Searcher's Preference Prediction Dongyuan Lu, Qiudan Li