Travel Recommendation via Author Topic Model Based Collaborative Filtering

Shuhui Jiang[1], Xueming Qian[1,*], Jialie Shen[2], and Tao Mei[3]

[1] SMILES LAB, Xi'an Jiaotong University, China
[2] Singapore Management University, Singapore
[3] Microsoft Research, USA
jsh.0531.smiling@stu.xjtu.edu.cn, qianxm@mail.xjtu.edu.cn,
jlshen@smu.edu.sg, tmei@miscrosoft.com

Abstract. While automatic travel recommendation has attracted a lot of attentions, the existing approaches generally suffer from different kinds of weaknesses. For example, sparsity problem can significantly degrade the performance of traditional collaborative filtering (CF). If a user only visits very few locations, accurate similar user identification becomes very challenging due to lack of sufficient information. Motivated by this concern, we propose an Author Topic Collaborative Filtering (ATCF) method to facilitate comprehensive Points of Interest (POIs) recommendation for social media users. In our approach, the topics about user preference (e.g., cultural, cityscape, or landmark) are extracted from the textual description of photos by author topic model instead of from GPS (geo-tag). Consequently, unlike CF based approaches, even without GPS records, similar users could still be identified accurately according to the similarity of users' topic preferences. In addition, ATCF doesn't predefine the category of travel topics. The category and user topic preference could be elicited simultaneously. Experiment results with a large test collection demonstrate various kinds of advantages of our approach.

Keywords: Multimedia, Travel Recommendation, Author Topic Model.

1 Introduction

In our daily lives, travel planning is always a tedious and difficult task. Gaining useful information from the fussy raw materials via manual analysis of travel guide website like IgoUgo (www.igougo.com) could be very time consuming, especially when travelers face a new city. Personalized travel recommendation techniques [1-9], [10-14], which can effectively integrate user preferences (e.g., cultural, cityscape or landscape), are gaining more and more attentions due to various potential applications in real world [11],[12].

Users' photos on social media record their travel history and much information about daily life. As shown in Fig. 1, a typical Flickr user's photo contains metadata like "User Id", "tags", "Taken data" "Latitude" and "Longitude".

* Corresponding author.

X. He et al. (Eds.): MMM 2015, Part II, LNCS 8936, pp. 392–402, 2015.

Fig. 1. Example of Flickr image information

Except the GPS trajectory information, the textual descriptions (such as tags and comments) that user gave when sharing the photos to social media networks (e.g. Flickr), is an important clue to infer user's latent interests [12-14]. For example, if a user visits a gym, the information about where he has gone can be identified and extracted from the GPS trajectory data. However, more detailed information about his interest such as "football" or "vocal concert" can be gained via visual analysis over the images and related tags.

Collaborative filtering (CF) is a well-known personalized travel recommendation approach [1]. However, it generally suffer from well-known "sparsity problem" in recommendation. Travel related data from real domain can be is very "sparse" and it makes accurate similar user identification very difficult if the user has only visited very few POIs. Recently, topic model (TM) learning method is introduced to solve the "sparsity problem" in travel recommendation [2]. Basic idea of TM is to infer users' travel topic preferences from the POIs that user has visited previously. Then the user preferred POIs can be recommended with similar topics. Usually the topic is determined by the naive category information from recommender system [2]. Unfortunately, for the community photo sharing websites like Flickr and Panoramio, it is difficult to define the category of travel topics due to lack of the accurate topic classification.

Motivated by these concerns, we develop an ATM based approach to model social users to carry out personalized travel recommendation. Due to data complexity, effective user travel topic preference mining with only textual description is challenging. Natural language models such as PLSA [10], LDA [4] and Author Topic Model (ATM) [3] are often utilized to cluster words to discover the latent topics that are combined to form documents in a corpus. LDA robustly discovers multinomial word distributions of these topics [4]. However, they cannot model authors and documents simultaneously. ATM directly annotates the user's interest with automatically divided semantic classes with respect to the distribution of the labels.

In this paper, we present an ATCF based personalized POI recommendation method by effectively extracting and integrating user travel topic preference from their tags of photo sets on social media. As illustrated in Fig 2, our ATCF based travel recommendation approach consists of two major functional modules - offline mining module and online recommendation module.

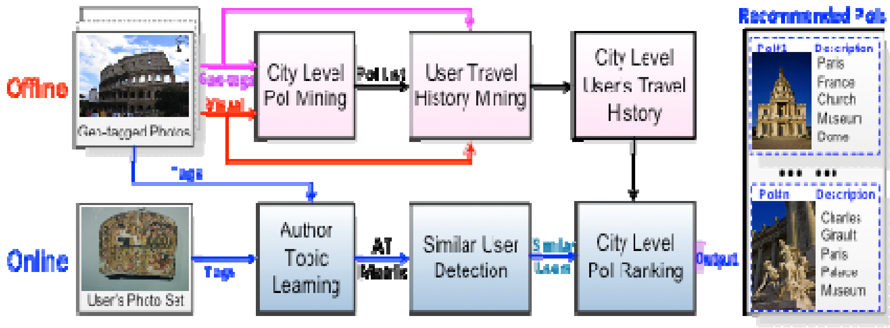


Fig. 2. A detail illustration of our travel recommendation system

In offline mining module, firstly, POIs of each city are mined using coarse-to-fine method from geo-tagged community-contributed photos by exploring both visual feature and geo-tags. Secondly, for each user, we mine each user's travel history (the POIs that user has visited) from geo-tags of user's community-contributed photos by a coarse-to-fine mapping method.

In online recommendation module, an ATM based approach is proposed to learn category of the topics and user travel topic preferences from the tags simultaneously. Secondly, similar users are mined based on the similarities of their topic preferences. Then, POIs in the new city are ranked based on the history of similar users' POIs visiting histories and the top ranked POIs are recommended to the user.

The main contributions of the research can be summarized as follows:

1) In this paper, we propose author topic collaborative filtering (ATCF) method based personalized travel recommendation systems. We utilize users' topic preferences as the law for collaborative filtering instead of location co-occurrences. It improves the sparsity problem of classical location based collaborative filtering (LCF).

2) We introduce an author topic model to adaptively elicit the topic category from tags associated with the Flickr photos. Using the scheme, topics about user preferences can be accurately extracted and applied to personalized travel recommendation. We also carry out large scale empirical study and the results show our approach enjoys great recommendation effectiveness.

The rest of paper is organized as follows: Section 2 is the introduction of related work. Section 3 and Section 4, the offline and online systems are described in detail. Section 5 presents experimental result and analysis. Finally, we conclude the paper in Section 6.

2 Related Work

In recent years, many different techniques have recently been developed to support travel recommendation based on different kinds of data. They include blogs [7], GPS trajectory [8], check in [5] and geo-tags [1,6]. Particularly, collaborative filtering algorithm shows its promising effectiveness in travel recommendation. The scheme is based on a Gaussian density estimation of co-occurrence space to cluster related geo-tags. They measure the similarity between two landmarks based on the similarity of travelers. By using the found similar user, a new trip plan can be made to a new location for a user [1]. While CF based recommendation methods demonstrate promising results, it suffers from the "data sparsity" problem. To solve this problem, topic model based methods are introduced to facilitate effective personalized travel recommendation [2, 9]. In [9], the authors conduct a study on exploiting online travel information by developing a tourist-area-season topic model. Bao et al., present a location-based and preference-aware recommender system that offers venues within a geospatial range [2]. They model each individual's personal preferences with a weighted category hierarchy using an iterative learning model in their offline system. However, in travel recommendation, it is difficult to find the authoritative category definition. Even though for check-in data, we could apply the original classification on website like Foursquare [2], for Flickr dataset with photos and textual descriptions [12-14] as shown in Fig.1, POIs are difficult to be categorized.

Distinguished from the existing POIs recommendation methods using CF, tags of photos on social media are used to represent user travel history in our system to mine user latent interest. Also, we use topic distribution to find similar users in a new city instead of the location co-occurrence. So accurate similar users could still be mined even the user has only visited very few POIs. Different from these mentioned topic model based methods, pre-defined the categories about travel topics are not required in our approach. By the author topic model, latent topics of travel could be mined adaptively.

3 Offline Mining Module

The offline mining module aims at mining POIs and all the users' travel history in the dataset from geo-tagged community-contributed photos. We propose coarse-to-fine POIs mining method when mining city-level POIs for each city. In "User Travel History Mining" part, we propose a coarse-to-fine-mapping method to mine user travel history for all the users using their geo-tagged photos.

3.1 City Level POIs Mining

The input is geo-tagged community-contributed photos with visual feature, tags and geo-tags. We have collected about 7 million social images from Flickr as Fig.1. In order to ensure that the noise photos of the dataset for each city are as less as possible, we use both the tags of city name and geo-tags of the location to double restrict the data of each city. After filtering, we get the geo-tagged photos of each city.

The input of **coarse-to-fine** method of POI mining is the geo-tag collection and visual feature collection of the city. First mean-shift clustering is used towards all the geo-tags of the photos in a city at a very small bandwidth as 0.0005, which is smaller than the radius of a landmark [6]. Each cluster contains a specified view. In this paper, we only use clusters containing at least 20 photos in the city and the number of users is no less than 10. Thus we get a list of clusters of a city denoted as *CL*.

We merge these clusters belonging to the same landmark (POI) by visual feature matching. First, we extract the 128D SIFT features for each image. Then we use bag of words (BoW) to present the SIFT descriptor. The size of the codebook is 61,944. Each image I_i is represented by BoW histogram. For a cluster C_n, the images belonging to it are represented by BoW histograms H_n, which is obtained by averaging the histograms of all the J images belonging to this cluster C_n. Their visual similarity $S(C_i, C_j)$ is measured by Euclidean Distance of H_i and H_j. For simplicity, we introduce a vector $AL=[A_1,A_2,\ldots,A_N]$ to record the assignment for the N clusters C_n to POI P_k as follows. If $S(C_i, C_j)$ is smaller than the threshold ThV, then the clusters C_i and C_j are belonging to the same landmark and we merge them together.

3.2 User Travel History Mining

This step aims at mining travel history for all the users. We make full use of the coarse layer clusters *CL* and assignment vector *AL* to get accurate travel history rather than directly compare the distances of a given image I_k to the centroids of the refined POIs *PL*, which is named as **coarse-to-fine mapping.**

Firstly, we determine the assignment of I_k to P_k by mapping user photos to cluster according to the geo-distances between user photos and the center of clusters. Then we determine which POIs the user has visited by mapping from *CL* to *PL* according to the assignment vector *AL*.

4 Online Recommendation

Online recommendation module aims at recommending POIs to a new user who has travel history about one city (city#a) and wants to visit a new city (city#b). We propose author topic learning based approach to mine user travel topic preference. Then POI recommendation is based on the history of similar user detection by the similarity of topic preference.

4.1 Author Topic Learning

Thus in this paper, we propose an ATM based approach to model social users to carry out personalized travel recommendation. The ATM is a generative model for document collections, which is able to extract information about authors and topics from large-scale text collections [3]. The input of this step contains two parts. The first part is a photo set I^u of user u_u with tag set τ_u. The second part is community users' photo sets with tags of each city. The output is topic preference distribution for each user.

In this section, first, the terminologies of ATM are introduced by combining the travel data. Then data processing of ATM procedure describes how to process the data as the input of ATM. At last the algorithm of ATM is shown.

4.1.1 Terminologies of ATM

In order to describe the model in our proposed method, we use the original terms (i.e., words, vocabulary, authors, and documents) to define the terminologies ATM in this paper as follows:

1) The **vocabulary** $V = \{1, 2, \ldots, N_d\}$ is the set of different tags of all the photos in a certain city. N_d is the size of V that represents the number of the different tags.

2) The **word** $w_i \in \{1, 2, \ldots, N_d\}$ represents the label of one tag of the photo, which can be considered as a representation of same tags of the photos. Note that each tag of an image is mapped to vocabulary V whose size is N_d through character matching and each tag can be represented by corresponding word w_i.

3) The document $d \in \{1, 2, \ldots, D\}$ corresponds to a tag set τ_j of the image I_j. So a user with NI images in the photo set has NI documents. So each photo with tags could be regard as one document.

4) The **authors** $a_d \in \{1, 2, \ldots, B\}$ is the label of the user who uploads the document d. $\{1, 2,\ldots,B\}$is the set of labels of the B users in the city. In our paper, each a_d has only one element, as each photo could only be uploaded by one user. a_d is only used ATM.

4.1.2 Data Processing for ATM

First, to construct the vocabulary V, we filter all the tags with both "stop words" and "Flickr-style words". A stop word can be identified as a word that has the same likehhood of occurring in those documents not relevant to a query as in those documents relevant to the query like "his", "on" and etc. "Flickr-style words" is a list of words frequently appear in Flickr tags but not in ordinary "stop words" like "Canon". We define these words manually after rank the words according to the frequency. After tag filtering, N_d tags without repetition construct the vocabulary V. Each tag in V has a label $w_i \in \{1, 2, \ldots, N_d\}$.

Second, for each user $a_d \in \{1, 2, \ldots, B\}$ who has upload document d (corresponds to an image), we map all the tags of the image to the V to get the label w. Thus all the tags of the city have been mapped to corresponding labels.

Third, we record the relationship between each document d with author ad. We also record the relationship between each word w with document.

4.1.3 Algorithm of ATM

The ATM is a Bayesian network as LDA. However, each author's interest is modeled with a mixture of topics by ATM, ATM is a hierarchical generative model in which each word w_i in a document d is associated with two latent variables, i.e., an author x_i and a topic z_i.

The generative process of ATM mainly consists of two steps: first an author x_i and a topic z_i are picked, and then, a word is generated according to the probability distributions. The details are as follows:

1) For each author $a_d \in \{1, \ldots, B\}$, choose a dimensional Dirichlet random variable θ_a~Dirichlet (α). For each topic $t \in \{1, \ldots, T\}$, choose $\phi_t \sim$ Dirichlet (β).

2) For each document $d \in \{1, 2, . . . , D\}$, given the vector of authors a_d, for each word w_i, indexed by $i \in \{1,\ldots,N_d\}$, do

(a) Conditioned on a_d, choose an author x_i~Uniform(a_d).
(b) Conditioned on x_i, choose a topic z_i~ Discrete(θ_{x_i}).
(c) Conditioned on z_i, choose a word w_i~ Discrete(ϕ_{z_i}).

As a result, we get *AT* matrix (Author Topic matrix) for all the users. *AT* is a sparse A x T matrix, where A is the number of authors, and T is the number of topics. AT (x,z) contains the times that a word w token associated with author x has been assigned to topic z.

4.2 POI Recommendation

In this section, we recommend POIs for a new user u_u in a new city#b according to his or her travel topic distribution AT_u mined by ATM in the city#a that has been visited before.

First, normalize each AT_i. Then we calculate the similarity between u_u and u_i from their author topic vectors using Cosine distance.

$$Sim(u,i) = \frac{AT_u \times AT_i'}{\| AT_u \| \bullet \| AT_i \|} \tag{1}$$

And then we rank the users in the city according to their AT distribution similarity. The top ranked *NS* users are selected as the set of similar users U^s.

Secondly, POIs in city#b are ranked according to the similar users' travel histories in city#b and then top ranked POIs are recommended to the user u_u.

5 Experimentation

We compare our proposed method (ATCF) with different approaches including recommendation by Popularity (PO), Collaborative Filtering (CF) and recommendation by LDA (LDA) to check the robustness of ATCF. The performance of these three methods and our proposed method are evaluated by criteria of MAP on Flickr dataset crawled by Flickr API on Flickr Website. MAP is one of the most well known criteria for measuring the relevance of recommendation. The descriptions of four compared methods are described as follows:

PO: First, POIs of the city are ranked according to how many users have uploaded photos of this POI.

CF: Location-based Collaborative Filtering is the most common way that can be most easily realized [1]. This baseline utilizes the users' location histories in a city to detect similar users.

LDA: To test the robustness of Author-Topic Model in the ATCF method, we replace the ATM with LDA model to mine user travel topic preference. Different from ATM, in LDA, we need to carry out an additional step to get user's topic distribution *AT*. In the first step, all tags of the photos in the city are allocated to different topics using LDA [4]. Each observed word is generated from a multinomial word distribution, specific to a particular topic. However in this step the relationship between authors and words, and authors and documents are not considered yet. Therefore, in the second step, we calculate the proportion of user's tags allocated to each topic mined in first step as *AT*. The other steps of LDA based approach are the same as ATCF.

5.1 Dataset

To facilitate comprehensive empirical study, we collected 7 million Flickr photos by open API. These photos are uploaded by 7,387 users and the heterogeneous metadata are associated with the photos.

After crawling Flickr photos, we only retain the photos with both tags and geo-tags from the original Flickr dataset. Though only tags are used to mine user's topic preference, geo-tags are also important to the recommendation system and evaluate experiments. On one hand, in offline system, geo-tags are involved in city level POIs mining and community users' travel history mining. On the other hand, in the evaluate experiments, the geo-tag that user labeled originally are regarded as ground truth of what the user have actually visited.

We select nine top popular cities to evaluate the performance of the five methods like that utilized in [1]. These nine selected cities are **Barcelona**, **Berlin**, **Chicago**, **London**, **Los Angeles**, **New York**, **Pairs**, **Rome** and **San Francisco**. We use the coarse-to-fine method in Part A of Section IV to mine POIs of these nine cities. Table II shows the corresponding number of users, POIs and photos in each city. There are 2,892 users, 307 POIs and 150,101 photos in total.

5.2 Performance Evaluation

In test data (Part A in this section), all the user photos retained to test contain both tags and geo-tags. And geo-tags, which record which POIs the user actually visited, are regarded as the **Ground Truth.** For example, if we recommend POIs in London to a user u_b, what he/she's travel history in London would be the Ground Truth. In the offline system, we mine travel history of u_b as $Q_b = \{q_1, ..., q_i, ...q_M\}$. To evaluate the performance, we compare the recommended POIs with POIs the user actually visited by his or her geo-tagged photos

We use MAP@n [1] to evaluate the performance of our method and the four comparative methods. It is one of the most well known criteria of the evaluation of recommendation system. In these two criteria, *n* denotes the number of POIs that we recommend to the user. We also provide the performance under MAP (without @n). In MAP, the number of recommended POIs is the same as the number that the user actually visited. The equation of MAP@n is as follows:

MAP@n: Mean average **precision** for a set of m users in the test data is the mean of the average precision scores for each user as follows:

$$MAP@n = (\sum_{i=1}^{m} AP_i) / m \qquad (2)$$

where AP_i is Average Precision of each user as follows

$$AP@n = (\sum_{i=1}^{n}\sum_{j=1}^{i} rel_j / i) / n \quad (3)$$

where rel_i is a relevance value. Suppose that we recommend n POIs in city#b to the user. To the i-th POI, we calculate how many POIs from 1-th to i-th POIs which we recommend are within the list of POIs user have actually visited in city#b. rel_i =1 if the user has actually visited the recommended POI, otherwise, rel_i =0. Then we average the results of n POI to get AP@n for the user.

5.3 Performance Comparison

Table 1 shows the recommendation results of ATCF (ours) on MAP in comparison with PO, CF and LDA. Similar users NS=40, and distance metric is Cosine distance. In LDA and ATCF, the number of topics is set to be K=50.

Table 1. Performance of POI recommendation on MAP of PO, CF, LDA and ATCF

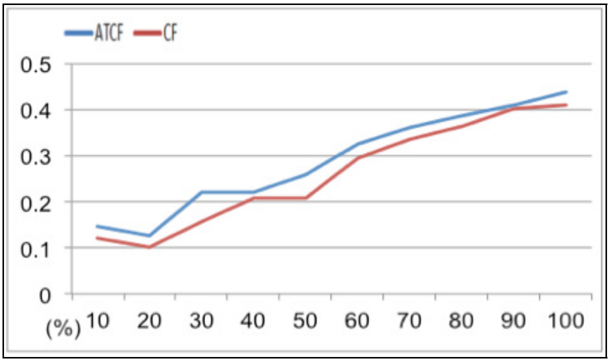
Perf.	PO	CF	LDA	ATCF
MAP	0.3408	0.4137	0.4166	**0.4225**
MAP@1	0.4861	0.5595	0.5678	**0.5876**
MAP@5	0.3557	0.4312	0.4361	**0.4483**
MAP@10	0.3076	0.4059	0.4005	**0.4115**
MAP@20	0.2642	0.3519	**0.3545**	**0.3545**
MAP@30	0.2438	0.3151	0.3163	**0.3184**

Table 1 shows the recommendation results of ATCF (ours) on MAP in comparison with PO, CF and LDA. Similar users NS=40, and distance metric is Cosine distance. In LDA and ATCF, the number of topic is set to be K=50.

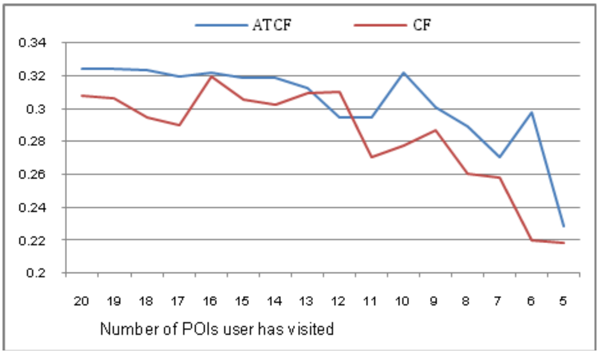
The performance on MAP of ATCF is 0.4225, which outperforms PO, CF and LDA by 8.17%, 0.88% and 0.59% respectively. Table 1 also shows the performance of MAP@n by n=1,5,10,20 and 30. We could see the performances of ATCF and LDA are higher than PO and CF. ATCF is the best when n=1,5,10 and 30.

5.4 Discussion

We conducted two experiments in order to evaluate the robustness of ATCF in "sparsity" condition. In the first experiment, we randomly sample the POIs from user travel history we mined in the baseline experiment. In Fig.3 (a), the x-coordinate means the proportion of POIs that we sampled. In the second experiment, we select the users whose travel histories are much sparser to construct the experimental data. In Fig.3 (b), the x-coordinate means users whose number of visited POIs is less than the certain value. In both experiments, NS=40, K=50.



(a)



(b)

Fig. 3. (a) MAP curves of CF and ATCF under sampled user travel history of different percentage. (b) MAP curves of CF and ATCF under users with sparser travel history.

The results of CF and ATCF under MAP criteria are shown in Fig.3 we could see that the performances of both CF and ATCF decrease when the data becomes sparser. Under the "sparsity" condition, the performance of ATCF is higher than CF in these two experiments. In the first experiment, the largest improvement of ATCF to CF is more than 0.5 when we sample user history at around 30%. In Fig. 3(b), in most cases, ATCF is 0.2 higher than CF. Only when the number of POIs is between 11 and 13, CF is higher than ATCF. When the number of POI is set to be 6, ATCF is 0.8 higher than CF.

6 Conclusion

This paper presents a novel author-topic collaborative filtering (ATCF) based personalized travel recommendation approach for social users. We model user travel preference and detect similar user simultaneously using author topic learning. User topic preference can be mined from the textual descriptions of photos by author-topic model (ATM) instead of history of locations from GPS (geo-tag) as most previous works.

However, there is still much work to be done. One of our future works is personalized travel route recommendation. We continue to crawl the photos from social media website. With more dataset, we could mine POI sequence instead of individual POIs.

Acknowledgments. This work is supported partly by NSFC No.61173109, No.61128007, No.60903121, Microsoft Research Asia, and Fundamental Research Funds for the Central Universities for Xueming Qian at Xi'an Jiaotong University. Jialie Shen was supported by Singapore Ministry of Education Academic Research Fund Tier 2 (MOE2013-T2-2-156), Singapore.

References

[1] Clements, M., Serdyukov, P., de Vries, A.P., Reinders, M.J.T.: Personalized Travel Recommendation based on Location Co-occurrence. CoRR, abs/1106.5213 (2011)
[2] Bao, J., Zheng, Y., Mokbel, M.F.: Location-based and Preference-Aware Recommendation Using Sparse Geo-Social Networking Data. In: ACM GIS 2012 (2012)
[3] Michal, R.Z., Griffiths, T., Steyvers, M., Smyth, P.: The author-topic model for authors and documents. In: Proceedings of Uncertainty in Artificial Intelligence, pp. 487–494 (2004)
[4] Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. The Journal of machine Learning research 3(5), 993–1022 (2003)
[5] Sang, J., Mei, T., Sun, J.-T., Xu, C., et al.: Probabilistic Sequential POIs Recommendation via Check-In Data. In: ACM SIGSPATIAL GIS, USA (November 2012)
[6] Cheng, A.-J., Chen, Y.-Y., Hang, Y.-T., Hsu, W.H., Mark Liao, H.-Y.: Personalized travel recommendation by mining people attributes from community-contributed photos. In: ACM MM 2011, NY, USA, pp. 83–92 (2011)
[7] Kori, H., Hattori, S., Tezuka, T., Tanaka, K.: Automatic Generation of Multimedia Tour Guide from Local Blogs. In: Cham, T.-J., Cai, J., Dorai, C., Rajan, D., Chua, T.-S., Chia, L.-T. (eds.) MMM 2007. LNCS, vol. 4351, pp. 690–699. Springer, Heidelberg (2006)
[8] Zheng, Y., Zhang, L.Z., Ma, Z.X., Xie, X., Ma, W.-Y.: Recommending Friends and Locations Based on Individual Location History. ACM Transactions on the Web 5(1), Article 5 (February 2011)
[9] Liu, Q., Ge, Y., Li, Z.M., Chen, E., Xing, H.: Personalized Travel Package Recommendation. In: Proceedings of 11th ICDM (2011)
[10] Hofmann, T.: Unsupervised learning by probabilistic latent semantic analysis. Machine Learning 42(1/2), 177–196 (2001)
[11] Shen, J., Cheng, Z., Shen, J., Mei, T., Gao, X.: The Evolution of Research on Multimedia Travel Guide Search and Recommender Systems. In: Gurrin, C., Hopfgartner, F., Hurst, W., Johansen, H., Lee, H., O'Connor, N. (eds.) MMM 2014, Part II. LNCS, vol. 8326, pp. 227–238. Springer, Heidelberg (2014)
[12] Qian, X., Feng, H., Zhao, G., Mei, T.: Personalized Recommendation Combining User Interest and Social Circle. IEEE Trans. Knowledge and Data Engineering 26(7), 1487–1502 (2014)
[13] Feng, H., Qian, X.: Mining User-Contributed Photos for Personalized Product Recommendation. Neurocomputing 129, 409–420 (2014)
[14] Qian, X., Liu, X., Zheng, C., Du, Y., Hou, X.: Tagging photos using users' vocabularies. Neurocomputing 111, 144–153 (2013)