

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

A New Approach to Travel Recommendation using Dynamic Topic Model and Matrix Factorization

Harshvardhan Hanmnatrao Patil, Prof. S.D. Satav

Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Hadapsar,
Pune, India

ABSTRACT: On web the rapid growth of online travel information have to face for tourists who have to choose from a large number of travel available packages to satisfy their personalized requirements to travel. The sparsity of user and location interactions makes it difficult to learn travel preferences, because a user usually visits only a limited number of travel locations. Static topic models can be used to solve the sparsity problem by considering user travel topics. However, all travel histories of a user are regarded as one document drawn from a set of static topics, ignoring the evolving of topics and travel preferences. In this paper, we propose a dynamic topic model (DTM) and matrix factorization (MF) based travel recommendation method. A DTM is used to obtain the temporally fine-grained topic distributions (i.e., implicit topic information) of users and locations. In addition, a large amount of explicit information is extracted from the metadata and visual contents of CCGPs, Check-ins, and POI categories datasets. The information is used to obtain user-user and location-location similarity information, which is a imposed as two regularization terms to constraint MF. User can view or get recommended places route on map. User's preference in mined through Lavenstine Distance Algorithm.

KEYWORDS: geo-tagged photos, check-in record, dynamic topic model, multimedia information retrieval, social media, Travel recommendation, travel recommendation.

I. INTRODUCTION

With the prevalence of intelligent cell phones and digital cameras with global positioning systems (GPS), people can casually take photos with geo-tags, and share these photos on social network sites (e.g., Panoramio and Flickr), which encourage users to upload the metadata information associated with photos . The number of community-contributed geotagged photos (CCGPs) has rocketed every day , and such large-scale CCGPs on the Internet are publically obtainable and cover most of the countries in the world . To alleviate the data sparseness problem in CCGPs data, additional information (e.g., category) is employed to obtain location similarity, which is then imposed as an extra regularization term of matrix factorization (MF) However, these methods ignore a large amount of important information (e.g., tourists' gender and age), which can be obtained and further exploited in similarity computation. Recently, static topic model (STM) is employed to model travel preferences by extracting travel topics from past traveling behaviors which can contribute to similar user identification. However, the travel preferences are not obtained accurately, because STM consider all travel histories of a user as one document drawn from a set of static topics, which ignores the evolutions of topics and travel preferences. To address these problems, in this paper, a dynamic topic model (DTM) and MF based method is proposed to recommend travel locations using ubiquitous data. In particular, a DTM is used to obtain the temporally fine-grained topic distributions (i.e., implicit topic information) of users and locations. A large amount of explicit information of users and locations are mined from the metadata and visual contents of CCGPs, Check-ins, and Point of Interest (POI) categories datasets. To address the sparsity of user-location

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

interactions, the information is used to obtain user-user and location-location similarity information, which is imposed as two regularization terms to constraint the factorization of user-location matrix. After that, the completed user-location matrix is obtained and each entry is used as a particular user's preference indicator of visiting a particular travel location.

II. MOTIVATION

The sparsity of user-location interactions makes it difficult to learn travel preferences, because a user usually visits only a limited number of travel locations. Static topic models can be used to solve the sparsity problem by considering user travel topics. However, all travel histories of a user are regarded as one document gets from a set of static topics, ignoring travel preferences.

III. REVIEW OF LITERATURE

1. We propose to conduct personalized travel recommendation by further considering specific user profiles or attributes (e.g., gender, age, race) as well as travel group types (e.g., family, friends, couple). Instead of mining photo logs only, we exploit the automatically detected people attributes and travel group types in the photo contents. By information-theoretic measures, we demonstrate that such detected user profiles are informative and effective for travel recommendation—especially providing a promising aspect for personalization. We effectively mine the demographics of individual and group travelers for different locations (or landmarks) and their travel paths. A probabilistic Bayesian learning framework which further mobile recommendation on the spot is introduced as well [1].

2. Present and compare 16 real life LBSNs, bringing into surface their advantages/ disadvantages, their special functionalities, and their impact in the mobile social Web. Moreover, we describe and compare extensively 43 state-of-the-art recommendation algorithms for LBSNs. We categorize these algorithms according to: personalization type, recommendation type, data factors/features, problem modeling methodology, and data representation. In addition to the above categorizations which cannot cover all algorithms in an integrated way, we also propose a hybrid k-partite graph taxonomy to categorize them based on the number of the involved k-partite graphs[2].

3. This paper presents a project called Knowing camera prototype System(KISS) for real-time places-of-interest(POI) recognition and annotation for Smartphone photos, with the availability of online geotagged images for POIs as our knowledge base. We propose a “Spatial+ Visual” (S+V) framework which consists of a probabilistic field-of-view(pFOV) model in the spatial phase and sparse coding similarity metric in the visual phase to recognize phone-captured POIs. Moreover, we put forward an offline Collaborative Salient Area (COSTAR) mining algorithm to detect common visual features (called Costars) among the noisy photos geotagged on each POI, thus to clean the geotagged image database. The mining result can be utilized to annotate the region-of-interest on the query image during the online query processing. Besides, this mining procedure also improves the efficiency and accuracy of the S+V framework [3].

4. Presents a personalized travel sequence recommendation from both travelogues and community-contributed photos and the heterogeneous metadata (e.g., tags, geo-location, and date taken) associated with these photos. Unlike most existing travel recommendation approaches, our approach is not only personalized to user's travel interest but also able to recommend a travel sequence rather than individual Points of Interest (POIs). Topical package space including representative tags, the distributions of cost, visiting time and visiting season of each topic, is mined to bridge the vocabulary gap between user travel preference and travel routes. We take advantage of the complementary of two kinds of social media: travelogue and community-contributed photos. We map both user's and routes' textual descriptions to the topical package space to get user topical package model and route topical package model (i.e., topical interest, cost, time and season). To recommend personalized POI sequence, first, famous routes are ranked according to the similarity between user package and route package. Then top ranked routes are further optimized by social similar users' travel records [4].

5. In this paper, we aim to study the semantics of point-of-interest (POI) by exploiting the abundant heterogeneous user generated content (UGC) from different social networks. Our idea is to explore the text descriptions, photos, user

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

check-in patterns, and venue context for location semantic similarity measurement. We argue that the venue semantics play an important role in user check-in behavior. Based on this argument, a unified POI recommendation algorithm is proposed by incorporating venue semantics as a regularized. In addition to deriving user preference based on user-venue check-in information, we place special emphasis on location semantic similarity [5].

6. In this paper, an author topic model-based collaborative filtering (ATCF) method is proposed to facilitate comprehensive points of interest (POIs) recommendations for social users. In our approach, user preference topics, such as cultural, cityscape, or landmark, are extracted from the geo-tag constrained textual description of photos via the author topic model instead of only from the geo-tags (GPS locations)[6].

7. In this paper we propose a location overview generation approach, which first mines location representative tags from travelogues and then uses such tags to retrieve web images. The learnt tags and retrieved images are finally presented via a novel user interface which provides an informative overview for a given location [7].

IV. SYSTEM OVERVIEW

Propose system, in this paper, a dynamic topic model (DTM) and MF based method is proposed to recommend travel locations using ubiquitous data. In particular, a DTM is used to obtain the temporally fine-grained topic distributions (i.e., implicit topic information) of users and locations. A large amount of explicit information of users and locations are mined from the metadata and visual contents of CCGPs, Check-ins, and Point of Interest (POI) categories datasets. To address the sparsity of user-location interactions, the information is used to obtain user-user and location-location similarity information, which is imposed as two regularization terms to constraint the factorization of user-location matrix. After that, the completed user-location matrix is obtained and each entry is used as a particular user's preference indicator of visiting a particular travel location. The main contributions of this paper are: (1) Present a DTM based way to obtain the temporally fine-grained topic distributions of users and locations. A large amount of explicit information is extracted from CCGPs, Check-ins, and POI categories datasets. (2) Present a MF based framework to address the sparsity of user-location interactions, by integrating the explicit and implicit information of users and locations as associated regularizations. Finally user gets recommended places map for travelling. User can view hotel, ATM on map.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

V. SYSTEM ARCHITECTURE

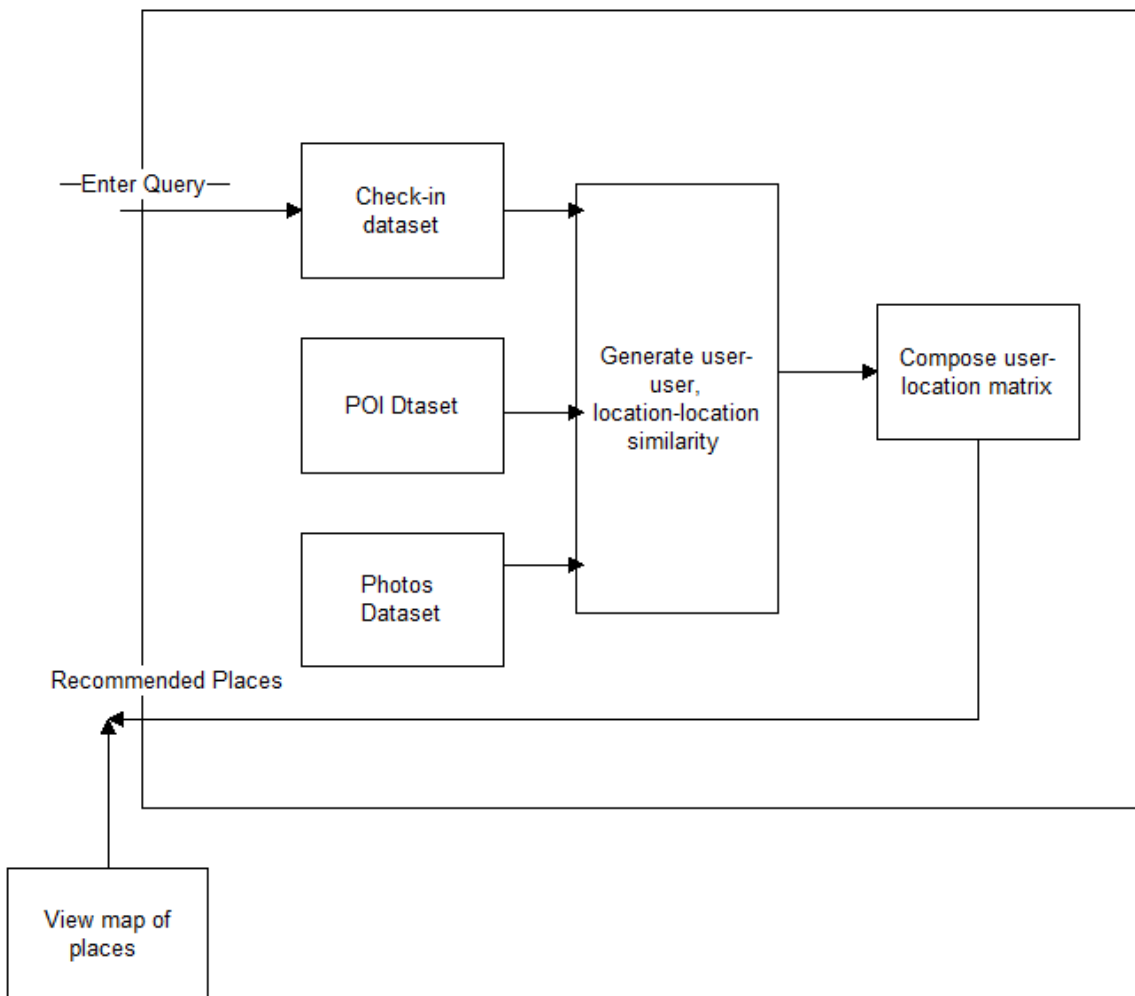


Fig. 01 System architecture

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

VI. MATHEMATICAL MODEL

Similarity between locations is calculated by Equations 1-3, and then the similarity matrix Mll is built, which is utilized to help the factorization of user-location matrix.

$$Sim(l_p, l_q) = \frac{\sum_{i=1}^{k'} f_{pi} \times f_{qi}}{\|l_p\| \cdot \|l_q\|}$$

$$\|l_p\| = \sqrt{\sum_{i=1}^{k'} f_{pi}^2}$$

$$\|l_q\| = \sqrt{\sum_{i=1}^{k'} f_{qi}^2}$$

where f_{pi} and f_{qi} represent the i th feature of locations l_p and l_q , respectively. ' k ' is the number of features.

$$Score(u_i, l_j) = U_i^T L_j$$

Above equation presents, preference score of user u_i to location l_j

VII. ALGORITHMS

1. Dynamic topic model:

The temporally fine-grained implicit features of users and locations are obtained by using DTM. A large number of explicit features of users and locations are extracted from the metadata and visual contents of CCGPs, Check-ins, and POI categories datasets. is utilized to build the profiles of users and locations. In this study, a plenty of features are explored to describe users and locations, which can gets diverse interesting information from the past traveling behaviors of users

Explicit Information:

Here we get following information:

1. Weather information: This information is helpful for constructing user and location profiles because of the fact that weather conditions can influence where a tourist wishes to visit.
2. Stay time information: This information is helpful for revealing tourist personal interest to a location.
3. Check-ins information: To obtain the fine-grained check-ins distribution of a travel location, we count the number of check-ins generated by users in POIs nearby the travel location in different time slots.
4. Gender and age information: Tourist gender is useful for describing a location. After the explicit information is extracted, we build the explicit feature vectors of users and locations.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

Implicit Information:

DTM Recommendation method work as:

- (1) The vocabulary V denotes a set of travel locations L' .
- (2) The word which represents one of the locations in L' .
- (3) The document N_t in A_t corresponds to the travel histories H_{it} of each user i in time slice t . Note that, the travel histories of all users are grouped by time slice t (e.g., by year).
- (4) The document set A_t corresponds to the travel histories D_t of all users in time slice t , i.e., $D_t = H_{1t}, H_{2t}, \dots, H_{mt}$. Through this model, we can obtain the topic distributions of users and locations in time slice t .

2. Matrix Factorization

User to get use- user similarity and location-location similarity . Here after we obtain explicit and implicit features, we get user and location profiles.

Location similarity matrix:

Base on location profiles, the similarity between locations is calculated by Equations 13, and then the similarity matrix M_l is built, which is utilized to help the factorization of user-location matrix. Each entry in M_l represents

the similarity between locations l_p and l_q . In M_l , location similarity $Sim(l_p, l_q)$ is ranging from 0 to 1, and a larger value means two locations are more similar. Location similarity matrix:

Similarly, the user-user similarity matrix M_u . Then combining user and location matrix we get user-location matrix.

3. 1.LEVENSHTTEIN DISTANCE ALGORITHM.

This is used to get POI mining from data which we get from Facebook at the Time of registration.

Input.

1. Text string $x = a_1$ for matching
2. Data from Facebook
3. Distance to match string x in Facebook data

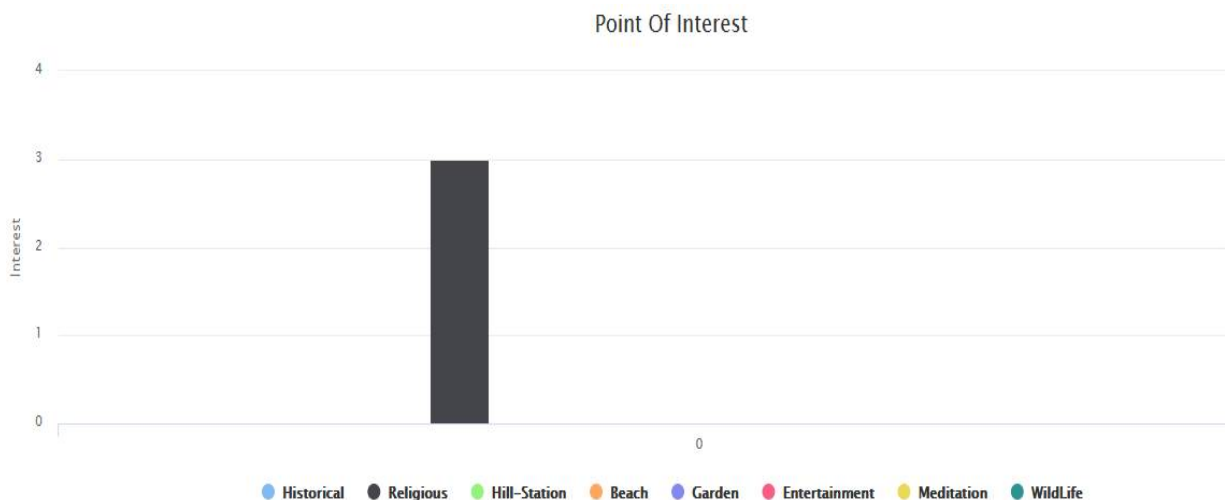
Processing: System will get POI as string and will match that string in Facebook data

According to specified distance.

Output: no. of time matched string with Facebook data

VIII.EXPERIMENT RESULT

1.Point of interest of each user:



Graph 1:POI Of each user

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

Explanation: Graph 01: shows point of interest of each user for each preference according to his like count .

2. Searching city of each season:

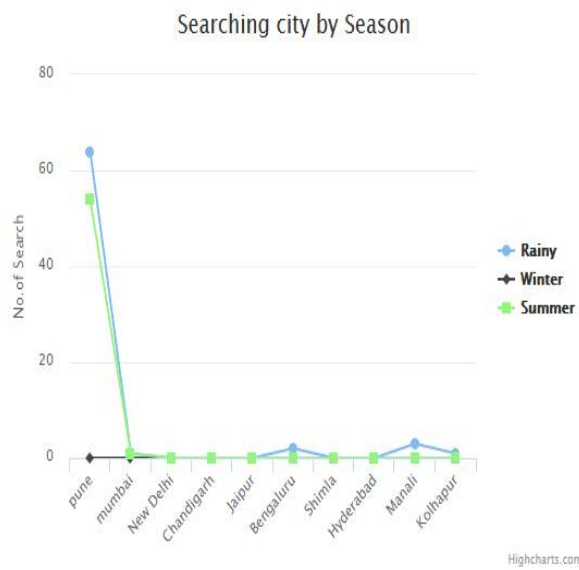


Fig 03: Searching city of each season:

Explanation: Graph Shows no. of searches of each city in particular season.

Result Table:

	summer	winter	Rainy
Pune	55	3	21
Mumbai	2	4	2
Chandigarh	2	2	3
Jaipur	3	1	2
bengaluru	3	4	1

Table 01. Shows no. of searches for particular city in different season for one month.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

IX. CONCLUSION

In proposed system we recommend the places to user according to his interest. for that we present a DTM and MF based method to recommend travel locations using ubiquitous data. The temporally fine-grained implicit features of users and locations are obtained by using DTM. A large number of explicit features of users and locations are extracted from the metadata and visual contents of CCGPs, Check-ins, and POI categories datasets. To address the sparsity of user-location interactions, these explicit and implicit features are used to obtain user-user and location-location similarity information, which is imposed as associated regularization terms to constraint the factorization of user-location matrix. Here user can view places on map like route. User can view ATM ,Hotel on map.

REFERENCES

- [1] Y. Y. Chen, A. J. Cheng, and W. H. Hsu, "Travel recommendation by mining people attributes and travel group types from community-contributed photos," *IEEE Transactions on Multimedia*, vol. 15, no. 6, pp. 1283-1295, Oct. 2013.
- [2] P. Kefalas, P. Symeonidis, and Y. Manolopoulos, "A graph-based taxonomy of recommendation algorithms and systems in LBSNs," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no 3, pp.604-622, Mar. 2016.
- [3] P. Peng, L. Shou, K. Chen, G. Chen, and S. Wu, "KISS: knowing camera prototype system for recognizing and annotating places-of-interest," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no 4, pp.994-1006, Apr. 2016.
- [4] Personalized Travel Sequence Recommendation on Multi-Source Big Social Media
Shuhui Jiang, Xueming Qian *, Member, IEEE, Tao Mei, Senior Member, IEEE and Yun Fu, Senior Member, IEEE
- [5] X. Wang, Y. L. Zhao, L. Nie, Y. Gao, W. Nie, Z. J. Zha, and T. S. Chua, "Semantic-based location recommendation with multimodal venue semantics," *IEEE Transactions on Multimedia*, vol. 17, no. 3, pp. 409-419, Mar. 2015.
- [6] S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, "Author topic model based collaborative filtering for personalized poi recommendation," *IEEE Transactions on Multimedia*, vol. 17, no. 6, pp. 907-918, 2015.
- [7] Q. Hao, R. Cai, X. Wang, J. Yang, Y. Pang, and L. Zhang, "Generating location overviews with images and tags by mining usergenerated travelogues," in Proceedings of the 17th ACM international conference on Multimedia. ACM, 2009, pp. 801-804.
- [8] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Time-aware point-of-interest recommendation," in *Proc. SIGIR*, 2013, pp. 363-372.
- [9] J. D. Zhang and C. Y. Chow, "Spatiotemporal sequential influence modeling for location recommendations: a gravity-based approach," *ACM Transactions on Intelligent Systems and Technology*, vol. 7, no. 1, pp. 11, Jan. 2015.
- [10] J. D. Zhang and C. Y. Chow, "Point-of-interest recommendations in location-based social networks," in *Proc. SIGSPATIAL*, 2016, pp. 26-33.