Wanderlust : A Personalised Travel Itinerary Recommender

Tarun Surender Reddy, Jenila Rakesh Sanghvi, Deval Paresh Vora, Pratik Kanani
[1-3]Students, [4]Assistant Professor
Computer Engineering,
D.J Sanghvi COE, Vile Parle (w), Mumbai.

Abstract— The project "Wanderlust: A Personalized Travel Itinerary Recommender" uses travelogues and community contributed photos from social media to generate a travel sequence for the user, along with this the metadata (e.g., tags, geographical location, and date taken) mined from the photos is also taken into account for a better response. The approach takes user's individual Point of Interest (e.g., museums, wildlife, historic sites, beaches, etc.) into consideration for travel recommendation rather than creating the itinerary based on only user's travel interest (e.g., New York, Las Vegas, Dubai, etc.) which most of the existing systems fail to provide. The project makes use of topical package space model which incorporates hash tags, travel cost spread, time of travel of each point of interest which is then mined to coalesce the gap between optimized travel routes and user's travel interest. The project uses two main types of social media viz. travelogue and community-contributed photos. For a personalized topical package model and an optimized route, mapping of user inputs with the mined data from social media and with routes to the topical package space is performed. Important and famous routes are given a rank according to the similarity between user package and route package in order to recommend a personalized point of interest sequence to the user. After processing, a route is selected from the top ranked routes to further optimize it according to socially similar users' travel records. Representative images with other user's viewpoint and comments and reviews of POIs are shown to offer a more detailed and wide-ranging impression. Our project provides recommendation on the basis of 60 thousand Flickr images uploaded by ample number of users and approximately 20,000 travelogues covering 450 travel POIs in five famous cities, and show its effectiveness.

Keywords— Travel recommendation, Geo-tagged photo, Representative images, Point Of Interest and Social media.

INTRODUCTION

Nowadays, people are interested in travelling and searching for different tourist location for travel planning according to their interests. Social media has come out as a medium to fulfil the continuous needs for automatic travel recommendation. It offers great opportunities to address many challenging problems, like GPS estimation and travel recommendation. Travelogue websites are basically blogs which offers rich descriptions about landmarks and also about the travelling experience which are written by users. These data are not only useful for determining POIs i.e. points of interest but also gives an opportunity to recommend personalized travel POIs and routes based on user's interest. Existing studies on travel recommendation use the different types of social media data, GPS trajectory, check-in-data, geo tag and blogs which are used for mining famous travel POIs and routes [2][4]. The existing system for travel recommendation has certain flaws due to which it cannot meet user's personal requirements. Personalized recommendation of travel system uses location-based collaborative filtering method in order to recommend the POIs and optimized routes by mining user's travel history. In this method, social similar users are mapped based on the location co-occurrence of previously visited POIs and then POIs are ranked according to the similar users travel history. There are two problems in automatic travel recommendation that needs to be discussed when compared with static existing travel recommendation approach. First, the recommended POIs should be personalized to user interest since different users may prefer different types of POIs [11][14]. Second, it is important to recommend a sequential travel route that is a sequence of POIs rather than individual POI. Existing system on travel recommendation typically comprises of two problems. The first problem, most of the travel recommendation system focuses only on user topical interest mining without considering other crucial attributes like consumption capability of the user. And for the second problem, existing systems focuses more on famous route mining rather than considering user travel interest [1]. To solve the above enlisted challenges, the new system proposes Topical Package Model method which automatically mines user travel interest from two types of social media data, different user-contributed photos and travelogues. For the first problem, it considers user's topical interest with the attribute like consumption capability and preference of visiting time of user. As it is difficult to measure the similarity directly between user and route, the proposed system uses topical package model which maps both user's and route's textual descriptions to the topical package model to get user package and route package using topical package space [1].

LITERATURE SURVEY

For travel recommendation system different system uses different types of data to mine user interested POIs, for this mining mainly four kinds of social media data are used that is GPS trajectory, check-in-data, geo tag and blogs for recommendation. User generated travel log provide rich information for a recommendation system [5][10].

Table 1.Comparative study of different technique

Author	Description	Merits	Limitation
S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei	This method makes possible comprehensive points of interest recommendations for social users. User preference topics are extracted from the geo-tag constrained textual descriptio n of photos via the author topic model. [2]	Without GPS records, similar users can still be mined accurately according to the similarity of users' topic preferences.	Data would be even sparser and noisier.
J. Li, X. Qian, Y. Y. Tang, L. Yang, and T. Mei,	It proposes a system of hierarchical structure to estimate the GPS location for animage wit h hierarchical global feature clustering and local feature refinement.[6]	Save computational costs and improve the GPS estimatio n accuracy.	Challengin g to estimate image GPS for photos.
J. Sang, T. Mei, T. J. Sun, S. Li, and C. Xu	It explores the potential of location- based service, which is to suggest a package of sequential activities related to user context and interest. [3]	Significance in recommending sequentia l activities.	Its not considering longer check-in sessions
H. Huang and G. Gartner	Collaborative filtering to mine GPS trajectorie s for providing Amazon-like POI recommendations.[4]	More accurate recommendation compare with simple location- based methods.	Considers only travel interest, doesnt take into account point of interest
H. Kori, S. Hattori, T. Tezuka, And K. Tanaka	It describe a system that extracts typical visitor's travel routes based on blog entries and that presents multimedia content relevan t to those routes. A sequential pattern mining method	Using desktop or a mobile device its act like tour guide.	Does not give options for the travel itinerary, it gives just one itinerary.

	is used for route extraction. [10]		

PROPOSED SOLUTION

To mine users travel interest, a Topical Package Model(TPM) has been used where the data is fetched from community-contributed photos and travelogues [12].

To address the existing first challenge, Our system considers not only user's topical interest but also the consumption capability and preference of visiting time and season. It is difficult to map the users point of interest(POI), so here the users point of interest(POI) will be input to the topical package space model , and the algorithm will return an optimized route for each day as well as a personalised itinerary

First, tags of user's photo set are mapped to topical package space to get user's topical interest distribution. It is difficult to get user's consumption capability directly from the textual descriptions of photos [13]. But the topics user interested in could somehow reflect these attributes. For example, if a user usually takes part in luxurious activities like Golf and Spas, he is more likely to be rich.

Our system combines user topical interest and the cost, time, season distribution of each topic to mine user's consumption capability, preferred visiting time and season. After user package mining, Our system ranks famous routes through measuring user package and routes package.

At last, Our system optimizes the top ranked routes through social similar users' travel records in this city. Social similar users are measured by the similarity of user packages.

Social Media Mining System Construction

In the first module development of the system for the evaluation of our proposed model and thus making the system construction module with social media mining system is done.

Our system implements Topical Package Space(TPS) algorithm and mines the users interest using 2 social media data sources namely : community-contributed photos, where the users who have visited a particular place share their reviews about it by posting pictures and describing their experience via captions or hastags, and Travelogeus, where the users just comment about their experience having visited that place, like the best season to visit, cost factor and so on.

The reasons for using the combination of social media are

1. Captions/tags with pictures have a lot of noisy data which makes it redundant , hence Travelogues which describes the location/place much better are used.
2. Each user has its own POI as well as cost factor which makes it difficult to compare it with tags and associated photos. (3) Although the visiting time potrayed by the social media may be accurate, but due to the large processing of photos it becomes difficult to come to one conclusion. [15][16]

Topical Package Space (TPS) Algorithm

- User topical package model (user package) is learnt from mapping the tags of user's photos to topical package space. It contains user topical interest distribution (U), user consumption capability (U), preferred travel time distribution (U) and preferred travel season distribution.
- This module introduces how to extract the user package, which contains user topical interest distribution, user consumption capability distribution, preferred travel time distribution and preferred travel season distribution.
- First user's topical interest is introduced by mapping user's tags to the topical package space. Then, how to get topical space mapping method is introduced.
- Then mapping og the textual description (tags) of user's community photos to the topical package space to present the user's travel preference of different topics is done, which is defined as user topical interest distribution [7]. It is assumed that if a user's tags appear frequently in one topic and less in others, the user has a higher interest towards this topic.
- Our system uses the cost distributions of the all the topics and distribution of user's topical interest to present a user's consumption capability. If a user usually takes part in luxurious activities like Golf and Spas, his consumption capability is very likely to be. If a user usually takes part in some cheap things, his consumption capability is likely to be low, and the system tends not to recommend him luxurious topics.

map showing the route to a particular destination. [6].

Algorithm 1: *SlackRoutes*

```
1.  Preprocessing phase: Cluster the POIs and construct the listOfClusters
2.  while listOfClusters is not empty do
3.      ClusterSet ← listOfClusters.pop
4.      RouteInit (ClusterSet)                    } Route initialization
5       iterations ← 0
6.      while iterations < maxIterations do
7.          Repeat
8.              Insert
9.          until no further insertion is feasible
10.         if currentSolution is the best found then
11.             bestFoundSolution ← currentSolution
12.             iterations ← 0                    } Route creation
13.         end if
14.         Shake
15          Iterations ← Iterations + 1
16.     end while
17.     remove all POIs visited in the currentSolution
18. end while
```

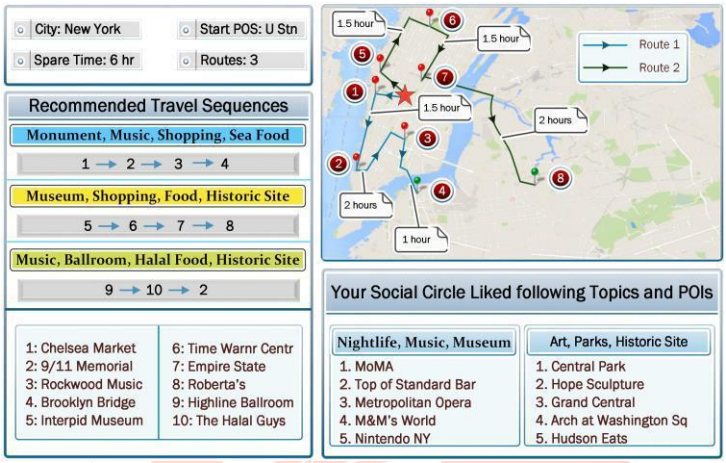

Figure 1. Clustering Algoritm.

Route Package Mining

Here the route topical package model takes as input the POI's of the user and travelogues data related to users POI, and it comes up with various optimized routes. Based on the best ranked route it then displays the user with an optimized route for each day of the itinerary.

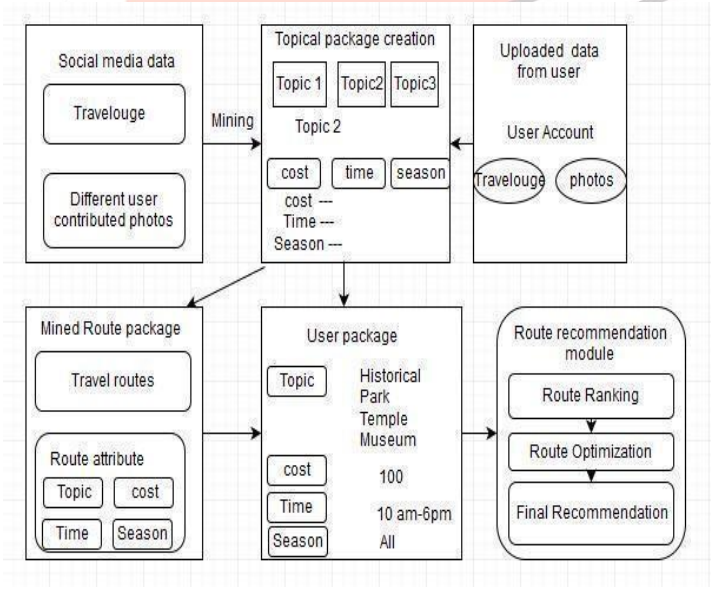

Figure 3. A visual example of travel routes

Here the ranking of the various routes are done based on a greedy algorithm where it searches for the best routes to reach a particular destination and displays the results based on a heuristic value.

After searching for the various routes, in order to display an optimized route kruskals algroithm is used, which helps us to find the shortest route, out of the various ranked routes, to reach a particular destination.

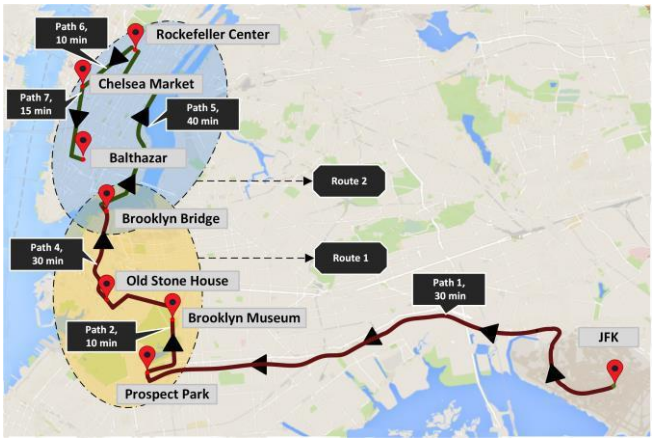

Figure 2. Travel pattern of a tourist who is interested in historical sites and nature.

Travel sequence recommendation

First step is to mine users POI with respect to travelogues and community-contributed photos. Once the mapping of the above said process is complete using topical package space algorithm, the system displays an optimized itinerary to the user, which ranges from day to day vists to a particular place for all age groups, to having a day freezed incase of an important meeting.
Next step is to display the most optimized route by making use of greedy and kruskals algorithm. This includes day to day optimized route as well as what distance to cover.

CONCLUSION

In this paper, Our system presents a personalized travel itinerary recommendation system by implementing topical package model using data mining form social media: travelogues and community-contributed photos. The advantages of our work are that firstly, the system automatically mines user 's topical preferences including the point of interest, cost and time and secondly the recommendation is not only providing point of interest but also travel sequence order, considering both the popularity and user's travel preferences at the same time. The system also provides user with flexibility to freeze a day or two for his/her personal work (e.g., meeting, conference, etc.) and successfully managed to show the travel itinerary and hotel bookings for comfortable stay in a single framework. Out project mines and ranks famous routes based on the similarity between user and route package and then optimizes the top ranked famous routes according to social similar users' travel records thereby providing user with the most efficient and feasible route.

FUTURE SCOPE

The current project gives user its own personalized travel itinerary based on his or her travel interests and point of interests along with hotel stay information. For future work, more type of data for mining user interest ca be used and also the system can provide new features which include providing air ticket details for a more convenient tour planning. Also a more detailed input can be taken from the user, asking the user about its eating preferences and based on that the system can suggest restaurants near every point of interests. Other miscellaneous things such as, giving the user specific privileges to tailor the itinerary by removing or replacing a particular place in the trip, can be added in the future. As the web-surfing era is about to end, the website can be converted into faster and easily accessible smartphone application and expand the project by providing itineraries for every place in the world. Also the website can be made more secured and different attacks are prevented using techniques like CAPTCHA [19], TextImage Ciphering [20] and Hybrid Key Distribution Systems [21].

REFERENCES

[1] Jiang, Shuhui et al. "Personalized Travel Sequence Recommendation on Multi-Source Big Social Media". IEEE Transactions on Big Data 2.1 (2016), pp.43-56, 2016
[2] S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, "Author topic model based collaborative filtering for
[3] personalized POI recommendation," IEEE Trans.Multimedia, vol. 17, no. 6, pp. 907–918, Jun. 2015.
[4] J. Sang, T. Mei, J.-T. Sun, C. Xu, and S. Li, "Probabilistic sequential POIs recommendation via check-in data," in Proc. 20th Int. Conf. Adv.Geographic Inf. Syst., 2012,402–405.
[5] H. Huang and G. Gartner, "Using trajectories for collaborative filtering–based POI recommendation," Int. J. Data Mining, Modelling Manage., vol. 6, no. 4, pp. 333– 346, 2014.
[6] J. Liu, Z. Huang, L. Chen, H. T. Shen, and Z. Yan, "Discovering areas of interest with geo-tagged images and check-ins," in Proc. 20th ACM Int. Conf. Multimedia, 2012,589–598.
[7] S. Jiang, X. Qian, J. Shen, and T. Mei, "Travel recommendation via author topic model based collaborative filtering," in Proc. 21st Int. Conf., MMM 2015, Sydney, N.S.W., Australia, 2015, pp. 392–402.
[8] Y. Ge, Q. Liu, H. Xiong, A. Tuzhilin, and J. Chen, "Cost-aware travel tour recommendation," in Proc. 17th ACM Int. Conf. Knowl. Discovery Data Mining, 2011, pp. 983–991
[9] Q. Yuan, G. Cong, and A. Sun, "Graph-based point- of-interest recommendation with geographical and
[10] temporal influences," in Proc. 23rd ACM Int. Conf. Inform. Knowl. Manage, 2014, pp. 659– 668.
[11] Y. Gao, J. Tang, R. Hong, Q. Dai, T. Chua, and R. Jain, "W2go: A travel guidance system by automatic landmark ranking," in Proc. Int.Conf. Multimedia, 2010, pp.123–132

[12] H. Kori, S. Hattori, T. Tezuka, and K. Tanaka, "Automatic generation of multimedia tour guide from local blogs," Adv. Multimedia Model.,vol. 4351, no. 1, pp. 690–699, 2006.
[13] Q. Liu, Y. Ge, Z. Li, E. Chen, and H. Xiong, "Personalized travel package recommendation," in Proc. IEEE 11th Int. Conf. Data Mining, 2011, pp. 407–416.
[14] X. Wang, M. Yu, L. Zhang, R. Cai, and W. Ma, "Argo: Intelligent advertising by mining a user's interest from his photo collections," in Proc. 3rd Int. Workshop Data MiningAudience Intell. Advertising, 2009, pp. 18– 26.
[15] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized recommendation combining user interest and social circle," IEEE
[16] Trans. Knowl. Data Eng., vol. 26, no. 7, pp. 1763–1777, Jul. 2014.
[17] Q. Hao, R. Cai, X. Wang, J. Yang, Y. Pang, and L. Zhang,"Generating location overviews with images and tags by mining user-generated travelogues," in Proc. 17th ACM Int. Conf.Multimedia, 2009, pp. 801–804
[18] V. Subramaniyaswamy, V. Vijayakumar, R. Logesh and V. Indragandhi, "Intelligent Travel Recommendation System by Mining Attributes from Community Contributed Photos", Procedia Computer Science, vol. 50, pp. 447-455, 2015.
[19] L. Pipanmekaporn and S. kamolsantiroj, "Mining Semantic Location History for Collaborative POI Recommendation in Online Social Networks", 2016 2nd International Conference on Open and Big Data (OBD), 2016
[20] Flickr dataset URL: https://api.flickr.com/services/rest/?method=flickr.galleries.getPhotos&api_key=ddf30709b27d0ec735670897e56f97d6&gallery_id=66911286-72157661097695568&format=json&nojsoncallback=1
[21] ViatorOnlineBookingAPI: "https://www.viator.com/tours/New-York-City/Statue-of- Liberty- and-Ellis-Island-Ticket/d687- 49200SOLEI?eap=makeMyTrip-9cbad44faebb487cbcdb40f6cb56cefc- 17895&aid=vba17895en".
[22] Y. Doshi, A. Sangani, P. Kanani, M. Padole, "An Insight into CAPTCHA", International Journal Of Advanced Studies In Computer Science And Engineering IJASCSE Volume 6, Issue 9, pp. 19-27. 2017.
[23] M. Padole, P. Kanani, "Textimage Ciphering", 2nd International Conference for Convergence in Technology (I2CT), 2017, pp - 926-930.
[24] P.Kanani, V.Kaul, K.Shah," Hybrid PKDS in 4G using Securred DCC", International Conference on Signal Propagation and Computer Technology (ICSPCT) 2014, pp- 323-328.