Travel Recommendation System Using Geotagged Photos

Akanksha Kumari
Ashish Singh
Rakesh Choudhary
Piyush Kumar Sarthi

Contents

Introduction

- Focus on developing a travel recommendation method that utilizes geo-tagged images in photo-sharing services.
- By sorting the photographs by their timestamps for each user, the geo-tagged photographs yield personal travel histories.
- The photographed geolocations are good recommendations in terms of finding attractive locations because the action of uploading a photograph can be regarded as a positive vote that the location is worth visiting.

Literature Survey

- Popescu et al. focused on the query with temporal constraints in terms of duration of the trip [1].
- Lu et al. proposed an interactive tourist recommendation approach which took into account a number of factors, e.g. duration of the trip and traveling cost, to help the tourist for trip planning [2],[3].
- Arase et al. focused on the detection of frequent trip patterns of people [4]. However, these works only consider the duration of the trip, traveling cost, and trip theme, not consider the context information of users when making recommendations for users.
- Context-aware travel recommendations take the context information of users into account, so these methods can provide more accurate recommendation results. Kurashima et al. incorporated present location information and preference of user into the probabilistic behavior model to make recommendations[5],[6].
- Cheng et al. proposed a probabilistic travel recommendation model which used people attributions for mobile travel recommendation and route planning [7],[8].
- Majid et al. considered weather context for travel recommendations of user [9],[10].

Outcome of Literature Survey

- When making recommendations, the previous works either do not take the similarity of users into consideration, or just employ a simple similarity computation.
- To counter the problems mentioned above, in our approach, we consider not only context information (e.g. season and weather) that influence the locations where users wish to visit, but also the similarities among users based on their traveling trips, using the modified longest common subsequence algorithm, which considers the order of travel locations of a user.
- This method can find the top n most similarity users and mine the preference of users.
- Also, to improve our results, we have also tried to include the events which are currently happening in a city and provide recommendations based on the number of people attending it and user interest.

Problem Definition

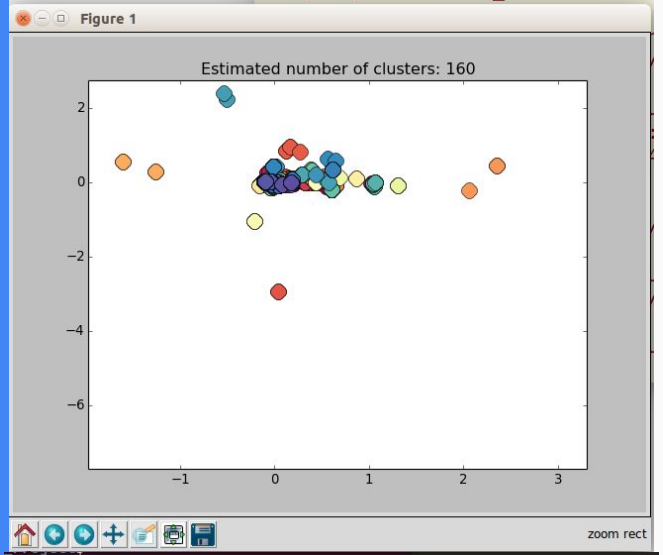
Propose a travel recommendation method based on the similarity of travel trips of users and the context history in which a location has been visited.

Objectives

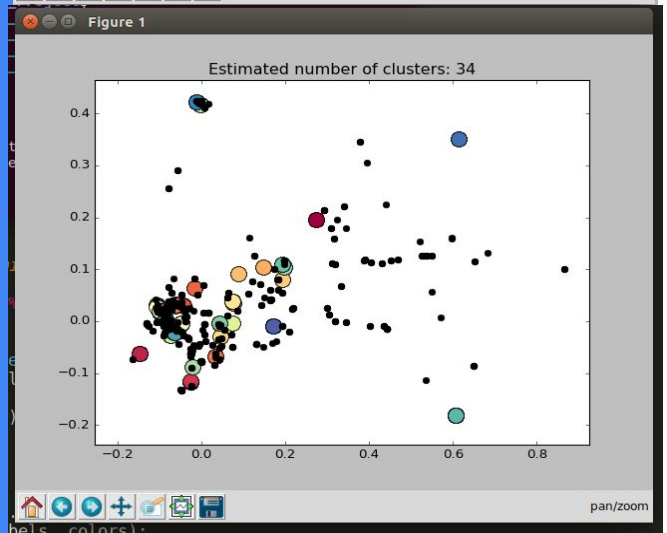
- Getting geo-tagged photos through various APIs like Flickr API.
- Refining the dataset through OpenRefine which would give us the tags, upload time, locations, coordinates etc. of the photos in the dataset obtained before.
- Data of the events/weather/location within the city that are currently happening in the cities in dataset has to be obtained via web scraping techniques. This data would be spanned over a period of 2-3 months.
- Clustering the locations within the city which the user wants to travel to depending on his preference. Experimentation with different clustering algorithms to identify the highly photographed location.
- We formulate the profiles of locations where profiling is mining of context and building a database.
- Modelling the preferences and similarities of users by building a user location matrix and a user-user similarity matrix.
- Recommendations based on the selected parameters (weather, season, events etc.)
- Comparison of our approach with the baseline methods like FR (Frequent Rank), CR (Classic Rank) and PCAR (Personalized Context Aware Rank).

Proposed Methodology

Results and Analysis



This figure is the result of implementation of SNN(Shared Nearest Neighbour) clustering algorithm. The number of clusters obtained was 160(in a very distributed manner) and the time taken to cluster the data was ~26 seconds(for 1517 filtered attributes).



This figure is the result of implementation of DBSCAN(Density-based Spatial Clustering of Applications with Noise) clustering algorithm. The number of clusters obtained was around 35(in a very precise and condensed manner) and the time taken to cluster the data was ~0.0063 seconds(for 1517 filtered attributes).

Results and Analysis

Profiling user location:

- We traverse through all the clusters except the last one(which is the noise cluster). We sort photos by descending order of timestamp for each user in a cluster. This is done by the following line of code:
  -->sorted_user_photos=indiv_cluster_data[indiv_cluster_data.User_ID==user].sort_values(by='Timestamp', ascending=False)

- Creating single_user_all_trips_list and all_user_all_trips_list :
  visit = visit.append(sorted_user_photos.iloc[photo_index+1])
  median_timestamp = visit.iloc[(len(visit)/2)].Timestamp
  all_users_all_trips_list[index].append(single_user_all_trips_list)

- Weather Context:
  After obtaining the median timestamp of a visit of a user to a particular location cluster, we obtain the weather context of that visit based on this timestamp obtained.We make use of Weather Underground API for this.
  url = 'http://api.wunderground.com/api/' + wunderground_api_key + '/history_' + str(date.date()).replace("-", "") + '/q/Maharashtra/Mumbai.json'

Results and Analysis

Building user-location matrix:

We can implement the matrix as:

```
Matrix = [[0 for x in range(no_of_users)] for y in range(no_of_locations)]
arr=np.where(unique_user_list1==user)
Matrix[arr[0]][index]=len(single_user_all_trips_list)
```

Building user-user matrix:

```
Matrix = [[0 for x in range(no_of_users)] for y in range(no_of_users)]
arr=np.where(unique_user_list1==user)
Matrix[arr[0]][index]=LCS( Travel_seq(user_i), Travel_seq(user_j))
```

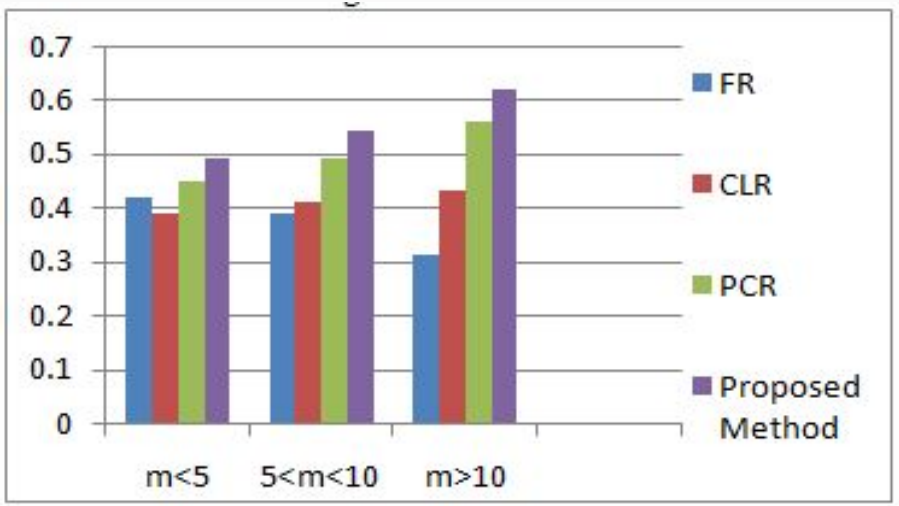
Recommendations

Making recommendations by collaborative filtering technique as described in the methodology.

Results and Analysis

Performance Comparison:

Method	FR	CLR	PCR	Proposed method
Precision(m<5)	0.42	0.39	0.45	0.49
Precision(5<m<10)	0.39	0.41	0.49	0.54
Precision(m>10)	0.31	0.43	0.56	0.62



Conclusion

Geo-tagged photos were collected through Flickr API and the dataset was refined through OpenRefine which gives us the tags,upload time,locations,coordinates etc of the photos in dataset obtained before. Clustering of the locations within the city is done, which the user wants to travel to depending on his preference.

Two different clustering algorithms were used, DBSCAN and SNN. Performance of DBSCAN is better than SNN.

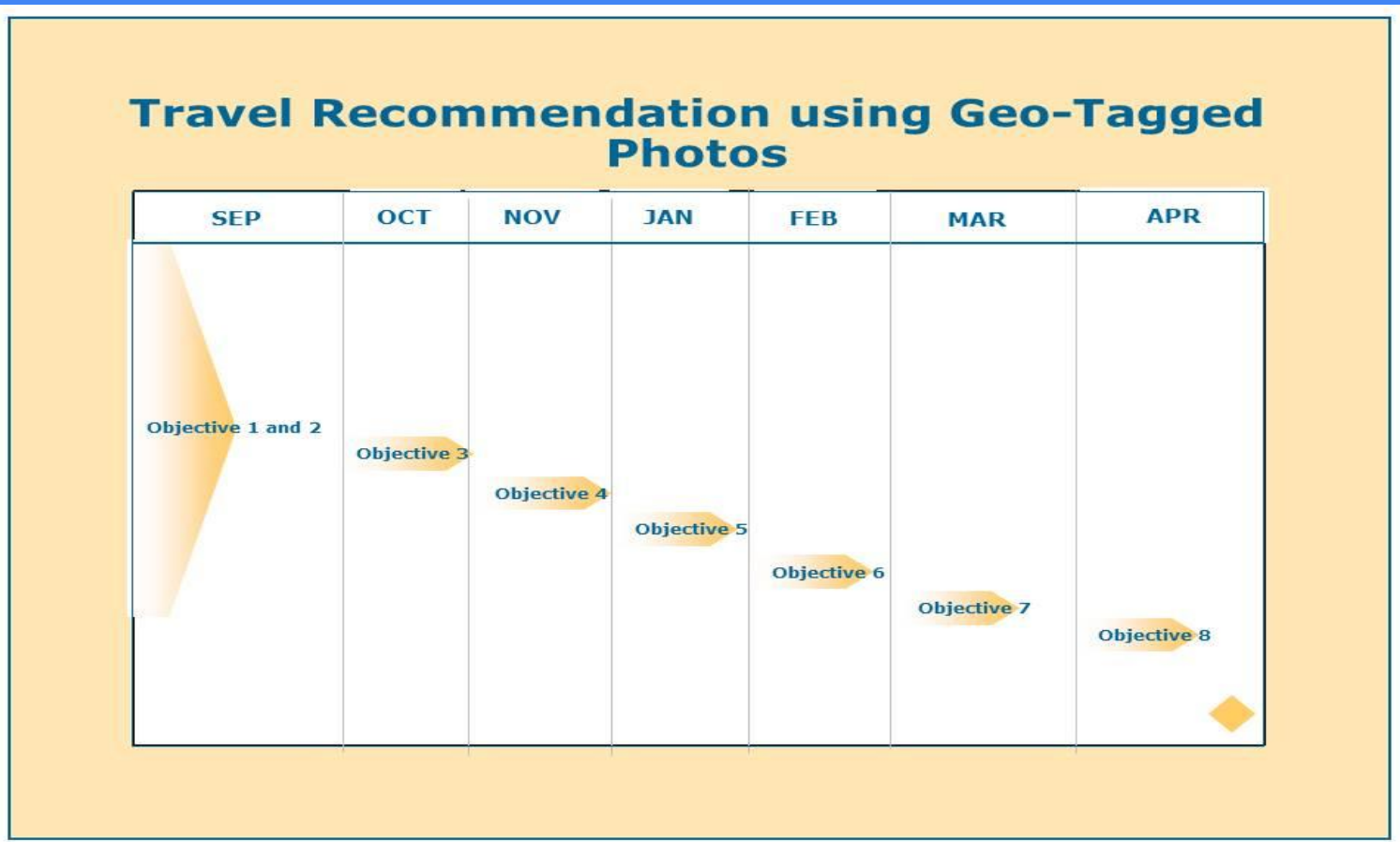
The clusters of locations were profiled with obtaining the weather information from the Wunderground API. Also, modeling of preferences was done by building user-location matrix and user- user similarity matrix.

Recommendations based on selected parameters(weather,season,live events etc.) were done. Comparison of our approach with the baseline methods like FR(Frequent Rank), CR(Classic Rank) and PCAR(Personalized Context Aware Rank) was done.

Future Work

In the future, attributes of photographers (e.g. gender and age) could be incorporated to further improve the accuracy of tourist preference mining.

Timeline


Travel Recommendation using Geo-Tagged Photos
SEP
OCT
NOV
JAN
FEB
MAR
APR
Objective 1 and 2
Objective 3
Objective 4
Objective 5
Objective 6
Objective 7
Objective 8

References

[1]A. Popescu, G. Grefenstette, “Deducing trip related information from flickr,” In Proceedings of the 19th International Conference on World Wide Web, ACM, 2009, pp. 1183-1184.

[2] H. Yin, X. Lu, C. Wang, N. Yu, and L. Zhang, “Photo2trip: an interactive trip planning system based on geo-tagged photos,” In Proceedings of the International Conference on Multimedia, ACM, 2010, pp. 1579-1582.

[3] X. Lu, C. Wang, J. M. Yang, Y. Pang, and L. Zhang, “Photo2trip: generating travel routes from geo-tagged photos for trip planning,” In Proceedings of the International Conference on Multimedia, ACM, 2010, pp. 143-152.

[4] Y. Arase, X. Xie, T. Hara, and S. Nishio, “Mining people's trips from large scale geo-tagged photos,” In Proceedings of the ACM International Conference on Information and Multimedia, ACM, 2010, pp.133-142.

[5]T. Kurashima, T. Iwata, G. Irie, and K. Fujimura, “Travel route recommendation using geotags in photo sharing sites,” In Proceedings of the 19th ACM International Conference on Information and Knowledge Management. ACM, 2010, pp. 579-588.

References

[6] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura, “Travel route recommendation using geotagged photos,” Knowledge and Information Systems, 2013, 37, pp. 37-60.

[7] A. J. Cheng, Y. Y. Chen, Y. T. Huang, W. H. Hsu, and H. Y. M. Liao, “Personalized travel recommendation by mining people attributes from community-contributed photos,” In Proceedings of the International Conference on Multimedia, ACM, 2011, pp. 83-92.

[8] 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, 2013, 15(6): pp. 1283-1295.

[9] A. Majid, L. Chen, G. Chen, H. T. Mirza, and I. Hussain, “GoThere: travel suggestions using geotagged photos,” In Proceedings of the 21st International Conference on World Wide Web., ACM, 2012, pp. 577-578.

[10] A. Majid, L. Chen, G. Chen, H. T. Mirza, I. Hussain, J. Woodward, “A context-aware personalized travel recommendation system based on geotagged social media data mining,” International Journal of Geographical Information Science, 2013, 27(4) pp. 662-684.

THANK YOU!