A Survey on Personalized Travel Recommendation System For Tour Planning

Mr. Wagh Rohan Nivrutti[1], Prof. N.G. Pardeshi[2]
[1]Dept of Computer Engineering
[2]Assistant Professor, Dept of Computer Engineering
[1, 2] SRES' COE Kopargaon, Maharashtra, India.

Abstract- *In this paper we present personalized travel sequence recommendation system that use travelogues, community contributed photos and the heterogeneous metadata (e.g.tags, geo-location, and date taken) associated with these photos to recommend user to plan his Travel Tour. We would construct a Topical package space by collecting travelogues,community contributed data.The system will analyse this synthetic data set to get the maximum Point of Interest between source to destination of user travel route. Unlike most existing travel recommendation approaches, our approach is not only personalized to users travel interest but also able to recommend a travel sequence rather than individual Points of Interest (POIs).Travelogues data is the data collected from travel sites (www.hellotravel.com) that offer rich depictions about historic points and voyaging background composed by clients. Moreover, group contributed photographs with metadata (e.g., labels, date taken, scope and so on.) via web-based networking media record client's day by day life and travel understanding. But can this information from various travelogue sites and group contributed photos used to recommend user a well planned travel sequence such that maximum POIs other than individual POIs would be covered the current system would fail here. Hence we proposed a novel approach to provide personalized travel sequence recommendation to users by analyzing his or her POIs. To obtain results such that maximum Point Of Interest would be covered we will be using Greedy algorithm so that we can get optimized results with maximum POIs.*

Keywords- Travel recommendation,Topical package space, Point Of Interests (POIs), Greedy Algorithm,Naive Bayes Algorithm.

I. INTRODUCTION

Presently automatic travel suggestion is a vital issue in both research and industrial area. Considering a real life scenario when a person plans a travel tour he or she has their Individual Point of Interest bonded within certain boundaries.In this case the person can miss some Point of Interest(POIs) other than Individual POIs between their routes i.e. between source to destination. Online social networking sites like Facebook, Flicker, Twitter and so forth offers incredible chances to address many testing issues, for example, GPS estimation and travel suggestion. Many travelogue sites (https://www.hellotravel.com) offer rich depictions about historic points and voyaging background composed by clients. Moreover, group contributed photographs with metadata (e.g., labels, date taken, scope and so on.) via web-based networking media record client's day by day life and travel understanding [1]. If we consider the current system and compare it with our system we would understand that the tourism sites are based on Package that the company offers to the clients which includes fixed tourist spots and fixed POIs. In such case the user may miss other interesting places that reside in between the source city and the destination city.Thus there was need of system that would analyze the data from Travelogues and data contributed from Social the Sites to suggest other POIs to User rather than his individual POIs. This helps users to plan his travel tour sequence such that maximum possible Point Of interests are visited that the user might be interested. In order to developed such system there was need of dataset i.e. data from Travelogues and Community Contributed Photos that includes (Longitudes, Latitudes,Season,Cost,Tags,POIs, Date Taken, Visiting time etc.). Thus to implement this system we should be build our own Tropical Package Space which constitute of some famous places in India. The Tropical Package space is combination of data from Travelogues and Community contributed data. Thus their felt need of developing system that would helps the user to plan his travel tours such that all the POIs that User might be interest gets visited.

II. LITERATURE SURVEY

Shuhui Jiang, Xueming Qian, [1] presented 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, the approach is not only personalized to users travel interest but also able to recommend a travel sequence rather than individual Points of

Interest (POIs). To recommend personalized POI sequence, first, they have ranked route according to the similarity between user package and route package and then displayed the optimized travel sequence such that it cover maximum POIs of user are covered.

S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, [2] they proposed an Author topic model based collaborative filtering (ATCF) method for personalized travel recommendations. Using author topic model users topic preference can be mined from the textual descriptions attached with users photos. Through author-topic model, travel topics, and a users topic preference can be elicited simultaneously. In this recommendation system, POIs are ranked according to similar users, who share similar travel topic preferences. This method overcomes the problem in location-based collaborative filtering, without GPS records, in author topic model based collaborative filtering method mine similar users accurately according to the similarity of user's topic preferences.

J. Sang, T. Mei, J.-T. Sun, C. Xu, and S. Li, [3] they explained the potential of location based service to overcome with an advanced recommendation problem activity plan, which is to suggest a package of sequential activities related to user context and interest. This type of recommendation system of point of interest is a probabilistic approach in which recommended POIs are relevant to user context i.e. current location, time, and check-in and personalized check-in history of the user. This approach of recommendation is highly motivated from a large-scale commercial mobile check-in data analysis, to rank a list of sequential POI categories and different POIs. The approach enables users to plan continuous activities going from one place to another.

Subramaniyaswamy and Vijayakumar, [4] they described the methods used to mine demographic information and provide travel recommendation to users. This paper also describes an algorithm adaboost to classify data and Bayesian Learning model for predicting desired location to a user based on his/her preferences.

III. SYSTEM OVERVIEW

The system we proposed is a personalized POI sequence recommendation system which could automatically mine user's travel attributes such as topical interest, consumption capability and preferred time and season. Travelogue websites (e.g. https://www.hellotravel.com) offer rich descriptions about landmarks and traveling experience written by user but can we use these data to recommend travel sequence for another User. A question arises that can we use this data of users who have already visited a certain places/spots to recommend another user to visit the place or not, further more we also have community-contributed photos with metadata (e.g., tags, date taken, Geo locations etc.) on social media record which gives user daily life and travel experience.Unlike traditional Travel recommendation system, the presented approach will not only recommend the User Personal Travel sequence but also will recommend a travel sequence rather than his/her Point Of Interest (POIs). The following *figure 3.1* gives architectural Overview of the system.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. The topical package space is combination of travelogue and community contributed data that we assumed to be collected from travelogues websites and social media sites. Thus 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)[1]. 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. Representative images with viewpoint and seasonal diversity of POIs are shown to offer a more comprehensive impression[1].

Thus the proposed system makes efficient use of the data collected to provide personalized travel sequence to the user in such a way that it contains maximum POIs that the user might be interested other than his/her individual POI. The system focuses on providing a personalized travel to user by analysing users POIs, Cost, Season, Tags etc.

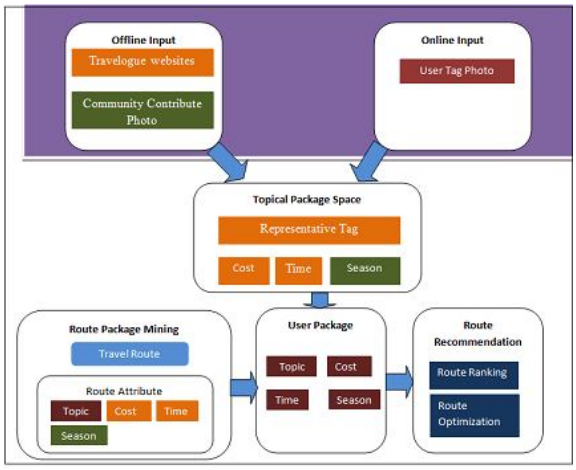

Figure 3.1: System Architecture

IV. MODULES

A. Topical Package Space

It includes construction of the synthetic topical package space in other words training dataset by the Combination of two social media: **Travelogues and Community-Contribute Data.**

- **Travelogue Data**

Travelogue data is the data collected from different traveling and Tourism Sites.

Following are the attributes associated with Travelogue Data.

1. **Departure Place**:
   It is the Source Place or City of the Travelogue Person.
2. **Visiting City**:
   It is the Destination City of the traveloguePerson.
3. **Topics**:
   Topics are the areas which person visited for e.g. Zoos, Forts, Temples, Hills, Rivers,Lakes,Museums etc.
4. **Multiple POIs:**
   It can be particular place or Spot for e.g. Sai Baba Temple, Juhu Beach etc.
5. **Tags**:
   It is the label attached to place for the purpose of identification other information for e.g Tag can be Temple, Beaches, and Forts etc.
6. **Visiting Time:**
   It gives best time the place can be visited.
7. **Cost:**
   It contains Budget required for the tour.
8. **Season**:
   It gives best season in which the place can be visited.

- **Community Contributed Data**

Community Contributed Data is the data from social networking site that describe user behaviour, andprovides more description about the location.

Following are the attributes associated with Travelogue Data.

1. **User id:**
   It is identity of user from Social Networking sites.
2. **Place or City:**
   It is the Place or City the User visited.
3. **Photos:**
   It contains the picture of the place which Community user visited.
4. **Tags:**
   It is the label attached to place for the purpose of Identification other information for e.g. Tag can be Temples, Beaches, Forts etc.
5. **Latitude and Longitude:**
   It provides location of the place or city.
6. **Season**:
   It is the season when the community user visited the place.

B. User Topical Package Space

It deals with user tour detail input for extraction of user package, which contains user topical interest Distribution, user consumption capability distribution, preferred travel time distribution and preferred travel Season distribution.

Following are the attributes associated with User Topical Package.

1. **Departure Place**:
   It is the Source city of the User.
2. **Visiting City**:
   It is the Destination city of the User.
3. **Topics Interest**:
   It includes the Topics which user is interested to visit e.g. Temples, Forts, Hill stations, Zoo, beaches etc.
4. **Time**:
   Preferred Time in which user wants to visit.
5. **Cost:**
   It is the Budget of User.
6. **Season**:
   The Season in User want visit place e.g.Spring, Summer, Winter etc.

C. Route Topical Package

Here the system performs travel route mining to get route topical package model.

- **Route Mining and Route Package Mining**

Here we mine POIs package including topical interest, Cost distribution, Time distribution and season distribution. Then to each route, we average all the POIs on the route to get route topical Package.

D. Travel Sequence Recommendation

- **Routes ranking**

Assume R =(r1, r2, ..rn) is a set of n travel routes mined. We rank these routes according to the similarity Between user package and routes packages. For user (uj) and route (ri), we measure the similarity of each attribute among topical interest(Ti),cost(C), Time(T) and Season(S) respectively.

To calculate topical interest(Ti) , cosine distance is applied to measure the similarity of distribution of users Topical interest and distribution of routes topics similar process is applied for rest attribute. Then (r1, r2,….rn) are ranked according to the order of topical interest from high to low. The ranked set of routes is denoted as R. If the route meets users interest, the score will be high, and it would be ranked at the top of the routes.

- **Route optimization**

After POI and route ranking module, we get a set of ranked routes R. Here, we further describe the optimization of top ranked routes according to social similar users travel records.

V. ALGORITHMS

A. Naive Bayes Algorithm

Naive Bayes is a collection of classification algorithms based on Bayes Theorem. It is not a single algorithm but a family of algorithms that all share a common principle, that every feature being classified is independent of the value of any other feature. Features, however, aren't always independent which is often seen as a shortcoming of the Naive Bayes algorithm and this is why it's labelled naive. Although it's a relatively simple idea, Naive Bayes can often outperform other more sophisticated algorithms and is extremely useful in common applications like spam detection and document classification.

As already said the basis of Naive Bayes algorithm is Bayes' theorem or alternatively known as Bayes' rule or Bayes' law. It gives us a method to calculate the conditional probability, i.e., the probability of an event based on previous knowledge available on the events.
More formally, *Bayes Theorem* is stated as the following equation:

$$P(AB) = P(B|A) * P(A) / P(B)$$

Where,

- P(A |B): Probability (conditional probability) of occurrence of event A given the event B is true
- P(A) and P(B): Probabilities of the occurrence of event A and B respectively.
- P(B|A): Probability of the occurrence of event B given the event A is true.

5.1.1 Implementation Steps for Algorithm

1. Input: S (Synthetic data Set.)
2. F:Convert the data set into a frequency table
3. T:Total Size of Dataset.
4. P: Calculate Prior Probability P (a)=F(a)/T
   (a attribute name)// Tag, City, Season, Cost.
5. Probability of Evidence E (e)=F(e)/T
   (e Evidence name)// Tag, City, Season, Cost .
6. L: Probability of Likelihood L (P(a))=F(e)//T(e)
   // Max attribute matching
7. While(size(F)==0)
8. R=higher (L(P(a)) //update Likelihood table
9. if R>R1 then update R
10. end while
11. end

5.1.2 Advantage of Naive Bayes Algorithm

- Relatively simple to understand and build.
- It is easily trained, even with a small dataset.
- It is fast.
- It is not sensitive to irrelevant features.

B. Greedy Algorithm

A greedy algorithm is an algorithmic paradigm that follows the problem solving heuristic of making the locally optimal choice at each stage with the hope of finding a global optimum. In many problems, a greedy strategy does not in general produce an optimal solution, but nonetheless a greedy heuristic may yield locally optimal solutions that approximate a global optimal solution in a reasonable time.

- **Implementation Steps for Algorithm**
  1. Start
  2. Select max Count of Solution(max POI achieve)
  3. If check given solution is equal to input then get final result and go to step 6
  4. Then check other max count of solution but is less than previous max count solution and go to step 2
  5. Then check given solution is equal to input then get final result go to step 6.(connected all POI & get final route)

6. Stop.

VI. SYSTEM REQUIREMENTS

A. Hardware Requirements

1. Processor : Intel Core i5
2. RAM : 2GB (min)
3. Hard Disk Space : 20Gb (min)

B. Software Requirements

- Operating System: Windows 7/8/10, Ubuntu 16.04.
- Coding Language : PHP
- Database : MySQL 5.0
- Netbeans 8.2,Note Pad++

VII. CONCLUSION

Thus the topic proposes the use and benefit of modern Personalized Travel Recommendation System that enable users to plan their travel sequence in best possible ways such that their maximum Topic of Interest would be covered. The system not only overcomes the drawbacks of current travel recommendation systems and tourism site but also provide user to plan their Personalized Travel Tour without been dependent on other sources. The systemwould also boost user travel experience in all possible ways.

VIII. ACKNOWLEDGEMENTS

"***Personalized Travel Recommendation System For Tour Planning***" has been a wonderful subject to research upon, which leads one's mind to explore new heights in the field of Computer Engineering. I dedicate all my dissertation works to my esteemed guide **Prof. N.G. Pardeshi**, whose interest and guidance helped me to complete the work successfully. This experience will always steer me to do my work perfectly and professionally. I also extend my gratitude to **Dr. D. B. Kshirsagar** (*H.O.D. Computer Engineering Department*) and **Prof. P. N. Kalavadekar** (*PG Co-ordinator*) who has provided facilities to explore the subject with more enthusiasm.

REFERENCES

[1] Shuhui Jiang, Xueming Qian, "Personalized Travel Sequence Recommendation on Multi-Source Big Social Media", Second InterIEEE TRANSACTIONS ON BIG DATA, VOL. 2, NO. 1, JANUARY-MARCH 2016.

[2] Shuhui Jiang, Xueming Qian,is, "Author Topic Model-Based Collaborative Filtering for Personalized POI Recommendations",ACM Computing Surveys , IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 17, NO. 6, JUNE 2015.

[3] Jitao Sang, Tao Mei, Jian-Tao Sun, Changsheng Xu, "Probabilistic Sequential POIs Recommendation via Check-In Data" National Lab of Pattern Recognition, Institute of Automation, CAS, Beijing.

[4] Subramaniyaswamy,Vijayakumar,Logesh R "Intelligent travel recommendation system by mining attributes from community contributed photos", 2nd International Symposium on Big Data and Cloud Computing 2015.

[5] Hongzhi Yin, Xiaofang Zhou, Bin Cui, Hao Wang, Kai Zheng "Adapting to User Interest Drift for POI Recommendation", IEEE Transactions on Knowledge and Data Engineering 2016.

[6] HUANG H., GARTNER G. "Using Trajectories For Collaborative Filtering Based POI Recommendation", Vienna University of Technology, VIENNA, AUSTRIA 2015.