IPHV8I3015X
International Journal Of Advanced Research and Innovation -Vol.8, Issue .III
ISSN Online: 2319 – 9253
Print: 2319 – 9245

CREATING TRAVEL ROUTE RECOMMENDATION SYSTEM BASED ON KEYWORD

#1T.SUCHARITHA, M.Tech Student,
#2G.SRILATHA, Assistant Professor,
#3Dr.M.SUJATHA, Associate Professor,
Department Of CSE,
JYOTHISHMATHI INSTITUTE OF TECHNOLOGICAL SCIENCES, KARIMNAGAR,T.S,INDIA.

ABSTRACT: In today's life recommender system plays the important role while user is on social networking site, online shopping website. For presenting the personalized recommendation according to the new user's demand is big challenge. We present an interactive visualisation tool for recommending travel trajectories. This system is based on new machine learning formulations and algorithms for the sequence recommendation problem. The system starts from a map-based overview, taking an interactive query as starting point. It then breaks down contributions from different geographical and user behaviour features, and those from individual points-of-interest versus pairs of consecutive points on a route. The system also supports detailed quantitative interrogation by comparing a large number of features for multiple points. Effective trajectory visualisations can potentially benefit a large cohort of online map users and assist their decision-making. More broadly, the design of this system can inform visualisations of other structured prediction tasks, such as for sequences or trees.

Keyword : *Route Visualisation, Travel Recommendation, Learning to rank, Map-Reduce.*

I. INTRODUCTION

The increasing no. of web services and increasing no. of internet users are growing rapidly. The data is collected from various sources like sensors, machine data, the data from the social networking, etc. The different varieties of evidence, the large volume of data and the velocity of data can cause the growth of data day by day. The Hadoop can process large volume of data very quickly. The recommendation system is the adaptive suggestion system to the user according to their interest. Here most of the recommender system can recommend the services to the user without analysing the previous reviews and current ratings. So the new user fails to meet the personalized service. The Keyword Aware Recommendation System (KASR) presenting the personalized recommendation based on the keyword[1]. Keyword service recommendation technique is presenting a personalized service recommendation list and recommending the most appropriate service to the users. Keywords are used to indicate user's preferences, and a User-based Collaborative Filtering algorithm is adopted to generate appropriate recommendations.

The KASR aims at calculating a personalized rating for each candidate service for a user, and then presenting a personalized service recommendation list and recommending the most appropriate services to him/her. The SmartCity recommendation system is having the list of the Smart Cities along with their services. The user can select the city and choose the service from the given list. This preference of the user is nothing but the keyword. Here keywords in the form of single keyword, multi-keyword or the sentence. This keyword can computes with previous users' similar preference from the database by using parsing and stemming techniques. If the previous reviews present in database relate to that keyword, then reviews are categorizing by using the sentiment analysis. The Map-Reduce is using for providing the most accurate and efficient recommendation to the user by computing the large set of reviews. The remaining paper is organized as follows: Section II describes the related work of various recommendation systems. Section III presents the proposed methodology for Smart-City Recommendation. In the section, IV states results and discussion and section, V presents conclusion and future work.

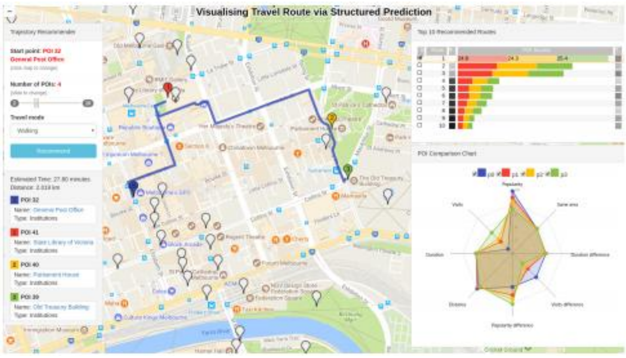

Figure 1: Travel route visualisation system1. Given a starting POI and the number of POIs to be visited, the system recommends multiple routes from travel history of tourists.

In this paper, we tackle the problem of sequence visualisation, specifically focussing on travel routes recommendation. A travel route is a sequence of POIs, and the sequence recommendation problem can be formulated as a structured prediction problem [3]. Based on a diverse set of features for individual and pairs of POIs, we train the prediction model with trajectory data extracted from geo tagged photos taken in Melbourne [1]. To visualise the suggested routes, we develop a novel tool that efficiently displays multiple suggested routes, which helps users understand the process behind the recommendations. Specifically, our system decomposes a total score of each route into a set of features and their corresponding scores, and shows the total score as a stacked bar plot of the features. The system also visualises the differences between POIs in a single route to show how POIs in that route can exhibit vast diversity. This visualisation helps tourists who want diverse experiences by choosing the best route among multiple recommendations. Generalising to a broader class of routes, such a visualisation could also help users of online mapping apps to make decisions on suggested travel routes, such as by trading off distance, traffic, and scenery.

Then can we find the most popular routes between two places and return them to the users? Ideally the routes recommended in this way should have improved quality since the popular routes are usually obtained by mining historical trajectories of drivers and can better indicate the driver's preference. But after further investigation, we have observed some problems of this approach. First, since users can specify two arbitrary places as source and destination, it happens that we cannot get enough support from historical trajectory data to obtain reliable results. Even worse, in areas where historical data are very sparse, the "popular" route can refer to a bad one with some random historical data. Therefore, a system that solely relies on those popular routes can perform badly at large scale. Second, there exist a number of popular route mining algorithms, each claiming some superiority from certain perspectives. When the results of these algorithms disagree with each other, it is still a pain for strangers to make a wise choice

II. RELATED WORK

Badrul Sarwar et. al. [2], here author concerns for the growth of business, the E- Commerce is using the group recommendation technique for providing better service to the user. Here they make the cluster of the user depending on their same types of reviews. The clustering of the population will divide the reviews of the user, and because of that, it gets affected by the accuracy of the service. Yan-Ying Chen et. Al. [3], here author works on the travelling recommendation it gives the user personal level recommendations for their trip. They keep track of the photos which they were clicked from various locations and depend on that they can give the recommendation to the user on the personal level They didn't consider the group level preferences for computation, and this is the limitation of the traveling system. Xin Cao et. al. [4], While searching the route from source to destination the user must have to give two specific keywords as source keyword and destination keyword. In the Google map by using these two keywords, it will recommend you the different route options. The user will choose the convenient route for his/her journey. J. Amaithi Singam et. al. [5],

In the optimal keyword search, the issues occurring when service recommender system implements in large data sets. It splits the services to the users and mainly focused keywords from the user's preferences. For generating keyword recommenders from the previous user preferences here using Hybrid Filter Algorithm. The result is shown here on Real-World datasets and reduces the processing time from large datasets. T.N. Chiranjeevi et. al. [6], PRS solving the challenges of the existing system like it provides the same result to the user's based on the evaluation and ranking or rating service. By using the collaborating filtering and Porter stemmer algorithm, it gives the suitable recommendation to the user. Here Personal Recommendation System is considered user's preference and necessity. This is using the Hashmap technique for faster keyword search for selecting correct reviews, and indexing method also used here in eliminating the articles like a, an, the, etc. X. Yang et. al. [7], The Bayesian-inference gives the recommendation for the social network. Here the user is get embedded with each other in the network. It is providing the accurate and personalized recommendation to the user. Each and everything is based on the network so the human being is facing difficulty for facing such interface. Xiaoyuan Su et. al. [8], In the survey of collaborative filtering techniques, there having the study of various techniques. There are three types of CF techniques such as Memory based technique, Model-Based Technique, Hybrid Based Technique. Each CF technique works in a different way. Yamini Nikam et. al. [9], In the survey on Service Recommendation Technique there are various recommender techniques such as Content-based, Knowledge Based, Social Network Based, Context Awareness based, Group Based, and Keyword Aware Service Based Recommendation. These recommendations are based on any one of the collaborative filtering technique. It is improving the scalability and the accuracy of the traditional systems. Khushboo R. Shrote et. al. [10], [11]

The sentimental analysis is known as the opinion mining analysis. It is using for analyse the users opining. This analysis is used for computing the positive and negative reviews from the user. And based on the analysis it provides the proper result. This survey is only about for how to apply sentimental analysis on the Keyword aware Service Recommendation for getting the appropriate result[12]. Rita Guimaraes et. al. [13], It presents a recommendation system

IPHV8I3015X
International Journal Of Advanced Research and Innovation -Vol.8, Issue .III
ISSN Online: 2319 – 9253
Print: 2319 – 9245

based on sentiment analysis of sentences extracted from Social Networks considering an algorithm which depends on the adverb found in the sentence. Depending on that adverbs the algorithm can performing the analysis. Here the recommendation system has low complexity and presents low perceived impacts on the analysis of energy consumption according to benchmark software. Kazuyoshi Yoshi et. al. [14], A hybrid music recommender system that ranks musical pieces by comprehensively considering collaborative and content-based data that is rating scores derived from users and acoustic features derived from audio signals. This mechanism is using his or her own preferences to select the music according to their choice. System accurately recommended pieces including nonrated ones from a wide variety of artists and maintained a high degree of accuracy even when new users and rating scores were added. Faustino Sanchez et. al. [15],

It presents a Recommender System for Sport Videos. Here the new recommendation method transparent to the user, who only has to consume videos as he or she would do in any video distribution platform. The system takes into account how the preferences of users change over time and based on that the recommendation was provided. This system is integrated on the client-side, and it avoids a lot of computational problem and congestion. Dr. M. Durairaj et. al. [16], In the news recommendation system it is using the content, knowledge and collaborative based techniques. The contents of the particular articles that a user has read in the past by analysing that with the help of content based approach. By using the collaborative filtering technique, the system makes the recommendation to the others, using known preferences of a group of the users. Shunmei Meng et. al. [1], here author gives details about Keyword Aware Service Recommendation, which is using for different types of domain thesaurus. And in KASR it will recommend the service by their previous and current reviews of the user for Hotel Reservation. Here keywords are used to indicate the user's preferences. The previous preferences of the user will be considered here for giving the appropriate service recommendation. It is using in Travelling recommendation, Hotel Reservation, Online Shopping, etc., and because of that, it is in the traditional way.

III. PROBLEM STATEMENT

In this section, we present some preliminary concepts and give an overview of the CrowdPlanner system. Table I summarized the major notations used in the rest of the paper.

TABLE I
SUMMARIZE OF NOTATIONS

Notation	Definition
R	a recommended route
$\mathbb{R}$	candidate set of recommended routes
p	a place in the space
l	a landmark in the space
$l.s$	significance of landmark l
$\mathbb{L}$	a landmarks set
$\mathbb{L}_{\mathbb{R}}$	the questioned landmark set of routes set $\mathbb{R}$
$d(l_i, l_j)$	Euclidean distance between landmarks l_i and l_j
w	a worker of the system
$\mathbb{W}_{\mathbb{R}}$	the selected workers of routes set $\mathbb{R}$

A. Preliminary Concepts Definition 1: Route: A route R is a continuous travelling path. We use a sequence [p1, p2, · · · , pn], which consists of a source, a destination, and a sequence of consecutive road intersections in-between, to represent a route. Definition 2: Landmark: A landmark is a geographical object in the space, which is stable and independent of the recommended routes. A landmark can be either a point (i.e., Point Of Interest), a line (i.e., street and high way) or a region (i.e., block and suburb) in the space. Definition 3: Landmark-based Route: A landmark-based route R¯ is a route represented as a finite sequence of landmarks, i.e., R¯ = {l1, l2, ..., ln}. In order to obtain the landmark-based route from a raw route, we employ our previous research results on anchor based trajectory calibration [21] to rewrite the continuous recommend routes into landmark-based routes, by treating landmarks as anchor points. Definition 4: Discriminative landmarks: A landmark set L is called discriminative to a set of landmark-based routes R¯ if for any two routes R¯ 1 and R¯ 2 of R¯, the joint sets R¯ 1 ∩L and R¯ 2 ∩ L are different. For example, L1 = {l3, l4} is discriminative to R1 = {l1, l2, l3} and R2 = {l1, l2, l4}, since the joint sets R1∩L1 = {l3} and R2 ∩ L1 = {l4} are different, but L2 = {l1, l2} is not discriminative to R1 and R2. Definition 5: Simplest Discriminative: An identifiable set L to R¯ is simplest discriminative to R¯ if removing any landmark from L, L is not discriminative to R¯ any more. Continuing with the previous example, L1 is not simplest identifiable to R¯, since after removing l3 from L1 it still identifiable to R¯, while sets L3 = {l3} and L4 = {l4} are.

B. Overview of Crowd Planner

Crowd Planner is a two-layer system (mobile client layer and server layer) which receives user's request from mobile client specifying the source and destination, processes the request on the server and finally returns the verified best routes to the user. Fig. 1 shows the overview of the proposed Crowd Planner system, which comprises two modules: traditional route recommendation (TR) and crowd-based route recommendation (CR).

The workflow of Crowd Planner is as follows: the TR module firstly processes user's request by trying to evaluating the quality of candidate routes obtained from external sources such as map services and historical trajectory mining; the CR module will generate a crowd sourcing task when the TR module cannot distinguish the quality of candidate routes, and return the best route based on the feedbacks of human workers of the system.

1) Traditional Route Recommendation Module: This module processes the user's request by generating a set of candidate routes from external sources (route generation component) and evaluating the quality of those routes automatically without involving human effort (route evaluation component). Control logic component: This component receives the user's request and controls the workflow of the entire system.

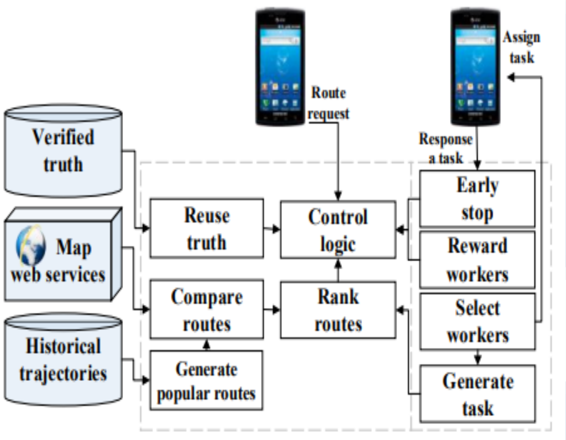

Fig. 1. System Overview

It also coordinates the interactions between the TR module and CR module. Once a user's request is received by the control logic component, it will invoke reuse truth component to match the request to the verified routes (truth) between two places at his departure time. If the new coming request is a hit of the truth, the system will return result immediately. Otherwise the component will invoke the route generation component to automatically generate some candidate routes and route evaluation component to evaluate the qualities of these candidate routes using the verified truth.

Route evaluation component: This component evaluates the routes using computer power and it provides an efficient way to reduce the cost of Crowd Planner, since it can largely reduce the amount of tasks generated. The component will firstly build up a candidate route set by invoking route generation component. If some of these routes agree with each other to a high degree, one of them will be selected as the best recommended route and added into a truth database with the corresponding time tag. If a best recommended route cannot be determined, the system will assign each candidate route a confidence score, which is generated by the verified truths and illustrates the possibility of the route to be the best recommended route. A route with the highest confidence score that is greater than a threshold η will be regarded to be the best recommended and returned to the user; otherwise the logic control will hand over the request to the CR module.

Route generation component: This component generates two types of candidate routes, the one provide by web services such as Google Map and the one generated from historical trajectories by using popular route mining algorithms, i.e., MPR, LDR and MFP.

2) Crowd Route Recommendation Module: Crowd route recommendation module will take over the route recommendation request when the traditional route recommendation module cannot provide the best route with high enough confidence. The module will generate a crowd sourcing task consisting of a series simple but informative binary questions (task generation component), and assign the task to a set of selected worker who are most suitable to answer these questions (worker selection component).

Task generation component: As the core of Crowd Planner, this component generates a task by proposing a series of questions for workers to answer. It is beneficial to have these questions as simple and compact as possible, since both the accuracy and economic effectiveness of the system can be improved. The design of this component will address two important issues: what to ask in questions and how to ask the questions. We will discuss the detailed mechanism of this part in Section III.

Worker selection component: This is another core component of Crowd Planner. In order to maximize the effectiveness of the system, we need to select a set of eligible workers who are most suitable to answer the questions in a given task, by estimating the worker's familiarity with the area of request. Technical details of this component will be presented in Section IV.

Early stop component: In most cases, we do not to need to wait for all the answers of the assigned workers. When partial feedbacks have been collected, this component will evaluate the confidence of the answer and return the result to the user as early as possible when the confidence is high enough.

Rewarding component: This component rewards the workers according to their workload and the quality of their answers. The reward points can be used later when they request a route recommendation in CrowdPlanner. In the following two sections, we will present the design and technical details of the two core components of CrowdPlanner: task generation and worker selection.

IV. TRAVEL ROUTE RECOMMENDATION

The travel route recommendation problem involves a set of POIs in a city. Given a trajectory query x = (s,l), comprising a start POI s and trip length l, the goal is to suggest one or more sequences of POIs that maximise some notion of utility. Following [3], we first cast travel recommendation as a structured prediction problem, which allows us to leverage

IPHV8I3015X
International Journal Of Advanced Research and Innovation -Vol.8, Issue .III
ISSN Online: 2319 – 9253
Print: 2319 – 9245

the well-studied literature of structured SVMs (SSVM) [6]. From a visualisation perspective, an advantage of the SSVM is the explicit representation of feature scores in its final decision process. Specifically, we can disassemble the final score of a route into feature scores of each POI and the transition between two adjacent POIs. We use hand-crafted POI features such as the category, popularity, and average visit duration of previous tourists. We also crafted transition features such as the distance and neighbourhood of two POIs to maximise the interpretability of the outcome.

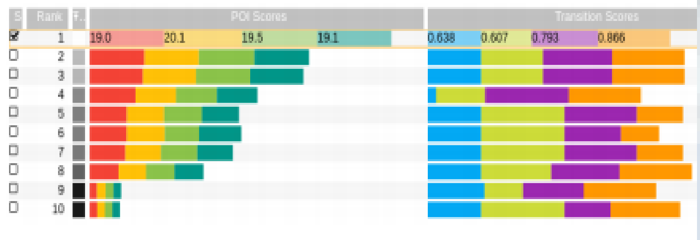

Figure 2: Visualisation of POI and transition scores for top 10 recommended routes. Each bar from left to right represents a relative score of each POI or transition along the route. The length of stacked bars represents the total score of the suggested route.

V. VISUALISATION

Our goal is to design an interactive visualisation system on top of the structured prediction framework. Figure 1 shows the overview of a live demo system, which consists of five major components: a map to display the suggested routes, an input box for user query (upper left), a stacked score of routes (upper right), a POI list box (lower left), and a radar chart to compare features of multiple POIs (lower right). The role and the construction of the four major components, besides the main map, are as follows:

Query input: A query consists of a starting POI and a trip length. Users can choose the starting POI by clicking icons on the map and can adjust the slide to set the trip length. In addition, three different travelling modes (e.g. bicycling, driving and walking) are supported, and we optimise the suggested routes for each mode.

Route score visualisation: The SSVM evaluates relevance scores of POIs and transitions in a candidate route to the given query and uses the sum of the relevance scores to determine the ranks of the routes. To visualise the POI and transition scores, we adopt a stacked bar representation [5], designed to support the visualisation of multi-attribute ranking. In Figure 2, the system decompose the scores of top 10 recommended routes into POI and transition scores via the stacked bar representation, where the size of each bar is proportional to the relevance score of the corresponding POI and transition in the route. Note that the POI and transition scores are scaled differently to support better visual discrimination2. For a seamless match between a route on the map and the corresponding POI scores in the bar plot, we use the same colour for both POI score and POI icon on the map. We also allow users to select multiple rows to visualise the corresponding routes on the map.

POI list: The POI list box provides the list of POI names and categories along the recommended route. The list is sorted according to the suggested visiting order, and again, the same POI colour is used to match the corresponding POI on the map. On top of the list, the system also provides an estimated travel time and total distance of the route. The POI list box is updated whenever a user selects a different route or the system makes a new recommendation. If more than one route is selected, the system displays the information of the most recent chosen route.

POI feature visualisation: We further provide a radar chart to analyse the variation between POIs in a single route. For example, in Figure 3, we compare two POIs (Melbourne Aquarium and Queen Victoria Market) in terms of POI features and their importance in the suggested route. The radar chart shows the corresponding POI feature scores when a user selects a route. In particular, the user can check/uncheck any POI in the selected route, and the feature scores of all checked POIs will be shown in the chart.

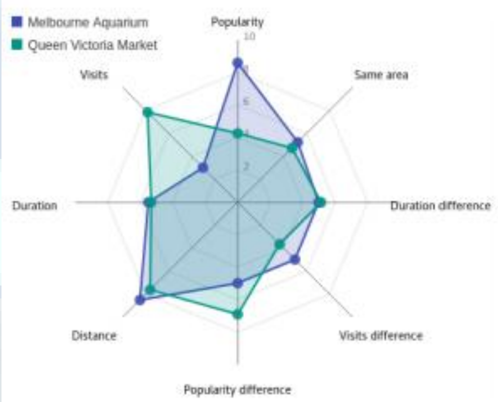

Figure 3: POI feature comparison between Melbourne Aquarium and Queen Victoria Market: the former scores higher on Popularity and Visits difference features whereas the latter scores higher on Visits and Popularity difference features.

VI. CONCLUSION

In this demonstration, we detail an interactive route analyser which helps the interaction between users and a route recommendation system. The system benefits from the explicit feature construction of the structured prediction model, and visualises recommended routes in terms of information on both the routes and the POIs.

We use many components to reduce the time cost and manpower cost of the route evaluating procedure, such as, truth reusing, landmark selecting and worker selecting, where in this paper we details the human evaluating related part. Landmark selecting automatically generates an identifiable set of landmarks with the highest mean significance. Worker selecting finds the top-k most eligible workers who have good knowledge of the task. Extensive experiments have been conducted involving a lot of volunteers and using a real trajectory dataset. We have demonstrated that the CrowdPlanner system can always give users the best routes. The MFP (Mining Frequent Path) has the highest possibility to give the best route. The ideas from this work open a new direction for future research, such as quality control of popular route mining algorithms, and mining latent factor which may affect drivers' driving routes.

REFERENCES

[1] Dawei Chen, Cheng Soon Ong, and Lexing Xie. 2016. Learning Points and Routes to Recommend Trajectories. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. ACM, 2227–2232.

[2] Dawei Chen, Cheng Soon Ong, and Lexing Xie. 2016. Learning Points and Routes to Recommend Trajectories (CIKM '16).

[3] Dawei Chen, Lexing Xie, Aditya Krishna Menon, and Cheng Soon Ong. 2017. Structured Recommendation. CoRR abs/1706.09067 (2017). https:arxiv.org/abs/ 1706.09067

[4] FAT/ML. 2017. Fairness, Accountability, and Transparency in Machine Learning. (2017). http://www.fatml.org, retrieved May 2017.

[5] Samuel Gratzl, Alexander Lex, Nils Gehlenborg, Hanspeter Pfister, and Marc Streit. 2013. Lineup: Visual analysis of multi-attribute rankings. IEEE transactions on visualization and computer graphics 19, 12 (2013), 2277–2286.

[6] Thorsten Joachims, Thomas Hofmann, Yisong Yue, and Chun-Nam Yu. 2009. Predicting structured objects with support vector machines. Commun. ACM 52, 11 (2009), 97–104.

[7] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. Computer 42, 8 (2009).

[8] Ching-Pei Lee and Chuan-bi Lin. 2014. Large-scale linear rankSVM. Neural computation 26, 4 (2014), 781–817.

[9] Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, and Shanika Karunasekera. 2015. Personalized tour recommendation based on user interests and points of interest visit durations (IJCAI '15).

[10] N.B. Robbins. 2005. Creating more effective graphs. Wiley-Interscience.

[11] Ben Shneiderman. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In Visual Languages, 1996. Proceedings., IEEE Symposium on. IEEE, 336–343.

[12] Wikipedia. 2017. Parallel coordinates. https://en.wikipedia.org/wiki/Parallel_ coordinates. (2017). Accessed: July 2017.