iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** V **Month of publication:** May 2018

DOI: http://doi.org/10.22214/ijraset.2018.5120

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue V, May 2018- Available at www.ijraset.com

Optimal Travel Route Discovery based on Topic Interest and Image Attributes

Asha Murali [1], Amina Beevi A [2]
[1] *M. Tech, Computer Science & Engineering, Sree Buddha College of Engineering, Kerala, India.*
[2] *Assistant Professor, Computer Science & Engineering, Sree Buddha College of Engineering, Kerala, India.*

Abstract: ***Recommendation in travel and tourism sector is emerging day by day. In order to suggest a travel route according to the user requirements (place, accomadation) is generated with the help of recommendation systems. The recommendation system is an emerging technique in many online sectors, in the case of travel and tourism has a major role for suggesting routes and other information. In travel route recommendation a system should consider user preference and travel history and possibility of route and so on. This is very huge task many algorithms are work behind this recommendation procedure to find and suggest a route or other information. Here user have very simple procedure like just providing inputs(source and destination) or provide a keyword for route search(temple, church, lake …)and then the system analyses the user query and generate requested output that satisfy user requirements. Here propose a travel route recommendation system to recommend route by considering the customer requirements. The route generation is basically done by two methods distance based route generation and MST based route generation. Here use kruskal as MST algorithm also providing a keyword searching facility. Experiments are taken by synthetically generated dataset for checking the efficiency of proposed system.***
Keywords: ***Minimum spanning tree, point of interest, travel route exploration, recommendation system, Global positioning system.***

I. INTRODUCTION

Data mining is defined as extracting information from a large amount of data. This is a computational process in a large data set that involves machine learning, statistics, and methods at the intersection of database systems. Apply intelligent methods to extract the basic process of data patterns. It is an interdisciplinary branch of computer science. The overall goal of the data mining process is to extract information from the data set and transform it into an understandable structure for further use. In addition to the original analysis steps, it involves database and data management, data preprocessing, model and reasoning considerations, interest metrics, complexity considerations, post-processing of discovery structures, visualization, and online updates. Data mining is the process of "knowledge discovery in the database" or KDD analysis.

In daily life, people are interested in travel plans that are interested in traveling and looking for different travel destinations. Social media has seen a constant demand for automatic travel recommendations. This has become an important issue in research and industry. Social media offers great opportunities for solving many challenging issues, such as GPS estimation and travel recommendations. The Travelogue website provides a rich description of landmarks and travel experiences written by users. These data can be used not only for reliable points of interest, ie points of interest, travel routes, but also to provide personalized travel interest points and route recommendations based on the user's interests. Current research on tourism recommendations uses different types of social media data, GPS trajectories, check-in data, geo-tags, and blogs for mining famous travel POIS and routes. The existing general tourism route planning system cannot meet the individual needs of users well. Personalization of the travel system recommends recommending POIs and routes by mining the user's travel history.

Location-based collaborative filtering is the best-known method. In this collaborative filtering method, socially similar users are mapped based on the location of previously visited POIs. POIs are then ranked according to similar user travel history. There are two problems with automatic travel recommendations when comparing existing travel recommendations. First, the proposed POI should be personalized according to the user's interests, because different users may prefer different types of POI. Second, it is important to suggest a continuous travel route, which is a series of points of interest rather than a single point of interest.The existing travel recommendation system does not solve these two problems well. The first problem is that most travel recommendations only focus on user topic interest mining and do not consider other attributes such as the user's spending power. For the second question, the existing research focuses more on the well-known route mining without considering the user's travel interest. In order to address the challenges of the order of travel locations and personalized recommendations for users, the new system proposes a partial package model approach that automatically mines user travel from two types of social media data,

©IJRASET: All Rights are Reserved

different user contributed photos and travel objects interest. For the first question, it considers the topical interests of the user's spending power and attributes such as user access time and season preference. It is difficult to directly measure the similarity between the user and the route. It is recommended that the system establish a theme package model and then map the text description of the user and the route to the theme package model to obtain the user theme package model (user package) and routing package. The model (route package) uses a partially encapsulated space. Compared with existing recommendation systems, this recommendation system is more suitable for travel plans.

II. RELATED WORKS

A. Y.-T. Wen, K.-J. Cho, W.-C. Peng, J. Yeo and S.-W. Hwang [2] proposed a keyword-aware travel route recommendation system. Given the huge number of registered data and photos in social media, it intends to discover travel experiences to promote travel plans. The system observes that when planning a trip, the user may have some keywords about his/her preferences in his/her travel. In addition, a variety of travel routes are also needed. In order to provide a variety of tourist routes, the system claims to extract features of more attractions (POIs). Therefore, proposed system is a keyword-aware Skyline Travel Path (KSTR) framework that uses knowledge extracted from historical flow information and users social activity in social sites. Explicitly, the system analyse the "where, when, who" problem by characterizing the geographical movement patterns, time effects, and social impact. Then a keyword capturing step (phase) is proposed to automatically classify the POI related tags into different types so as to effectively match the query keywords. The route reconstruction algorithm is further designed to build a route candidate that matches the user provided input. Provide diversified search results and explore the Skyline concept to arrange routes. In order to find the accuracy of given algorithm, extensive experiments were conducted on location-based real social network datasets.

B. T. Kurashima, T. Iwata, G. Irie, and K. Fujimura [3] suggest using geo-tags to recommend travel routes in photo-sharing sites. This article discusses the capability of the system to create geo-tagged photos that enable people to share the personal experiences as visitors at specific locations and times. Assume that the collection of geo-tagging photos for each photographer is a series of visited locations. The photo sharing website is an important source for collecting the location history of tourists. With the help of location information system can easily find out the landmarks associated with the travel route. Here the system contain a routes proposed method based on photographer history mastered by Flickr is proposed. The recommendation is performed by the photographer's behavioral model, which count the speciality of photographer’s landmark in each place. The system combines user preferences and current location information into a probabilistic behavior model by combining topic models and Markov models. The system demonstrates the effectiveness of the proposed method using a real data set that holds information from 71,718 photographers from the United States about the accuracy of the prediction of travel behavior.

C. Yukiko Kawai, Jianwei Zhang Hiroshi Kawasaki [4] proposed a travel propose a system by considering both social information and GIS. This article has conducted in-depth research on information recommendation and filtering technology. The traditional travel recommendation system can be considered as one of the information recommendation systems. Usually, the shortest path of time or distance can be calculated. Recently, travel recommendation systems for more general purposes have become an important research topic. System proposes an efficient travel route search system, which not only recommends a simple route to connect multiple tourist attractions, but also recommends a beautiful road. The system focuses on the visibility of scenic spots between one attraction and another, this is a major fact in selecting driving routes, but is not considered in the available way of tourisms proposals. To automatically retrieve tourist attractions, use Web information to propose personalized tourist attractions recommendation technology. Although in some areas, databases of famous attractions already exist and have been published, these areas are limited and often outdated. The proposed method automatically extracts points from the Web, so the system is versatile and up-to-date for large areas. To find a route with a fascinating landscape, the system calculates the route score based on the visibility of the scenic spot. After generating a candidate route using a GIS, a 3D virtual space is constructed and the Z-Buffer method is used to determine the visibility of the scenic spots for each candidate route. The system generate the guidelines and measure the ability of system.

III. SYSTEM OVERVIEW

A. Existing System

Existing works are scored and each route will be scored based on its characteristics (for example, the number of places, the acceptance of the residence). In general, the query results will have similar directions. The existing system aims to search for more types of routes based on the travel factors considered. Since high-ranking courses are often too similar to each other, this work considers the diversity of results by using Skyline queries. In the existing system, a system called the Photo2Trip system was

©IJRASET: All Rights are Reserved

developed that mixes a sequence of travel features including duration, season, user preferences, destination type and popularity to recommend travel itineraries. Another system arranges the route through location attraction, proper access time and distance to the query location.

1) *Drawbacks of Existing Syste*
2) query results of existing travel route recommendation services usually rank the routes simply by the popularity or the number of uploads of routes
3) The existing works focused on the efficient way to search for existing routes that cover all the pre-defined keywords
4) The existing system, restricting users to limited query options such as locations, activities or time periods.

B. Proposed System

When planning a trip, users always have a particular preference for travel. The system does not limit the user's limited query options (such as location, activity, or time period), but instead uses any textual description as a key to personalize requirements. In addition, a diverse and representative recommended travel route is also needed. The previous work elaborated on excavating and arranging existing routes from check-in data. In order to meet the needs of automotive travel organizations, the system claims that more POIs should be extracted. Therefore, proposes an efficient keyword-based representative travel route framework that takes advantage of the user's historical flow record and knowledge in social interactions [1]. Explicitly, the system designs a keyword extraction module to classify POI-related tags in order to effectively match query keywords.

The system further designs route reconstruction algorithms to build route candidates that meet the requirements [1]. In order to provide appropriate search results, here existing routes are taken in account to generate new route. In order to evaluate the effectiveness and efficiency of the proposed algorithm, extensive experiments were conducted on location-based real social network datasets.

The minimum spanning tree algorithm improves the recommendation strategy by considering the cost of the node. The cost is a composite score obtained from the POI score and the check-in score. The execution speed here has been improved. The kruskal algorithm is used here as MST. It found the least weighted edge connecting any two trees in the forest. It is a greedy algorithm in graph theory because it finds the minimum spanning tree that connects weighted graphs, increasing the arc cost per step. This means that it finds a subset of the edges, which forms a tree that includes every vertex with the smallest total weight of all the edges in the tree.

Image attributes come from uploaded images. Image attributes include LBP features, grayscale values, intensity, and finally a histogram. An image histogram is a histogram that is used as a graphical representation of the distribution of tones in a digital image. It plots the number of pixels for each tone value. Comparing the histograms of each blog's image and comparing them to each other helps to recommend routes based on image similarity.

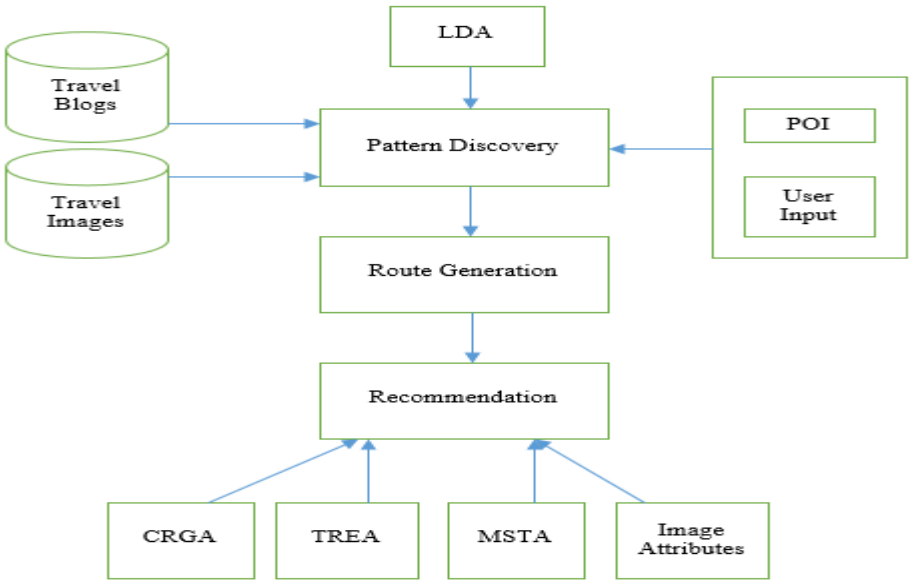


Fig 1. System Architecture

©IJRASET: All Rights are Reserved

IV. SYSTEM DESIGN

A. Modules

Optimal travel route discovery based on topic interest and image attributes consist of two modules.

1) *Admin Module:* The main part of the project is the admin. Which controls the entire system process. Executives do the main job of the project, such as the main concept of the travel route recommended by the administrator. In which admin set the details of the location, including the location name, description, longitude and latitude of each place, this helps recommend. The admin also set a travel package that does not include days, food and taste details, and total cost. Admin can also set comfort map administrators here to set up important places near each place (police station, train station, etc.). Next is interest modeling, where the admin can briefly review the date the blog was written by the user when the date was provided, write the blog on that date, and then tailor the blog to delete the is, was, etc. Stop words and then execute the interest model to find important words in the blog, which helps with recommendations. There are two main technologies used for this purpose here, such as MALLET and LDA algorithms.Then admin can also have the ability to find image similarity. This task is very helpful for recommending one user to another user based on blog image similarity. Admin process the blog images of each user and then find the similarity among all the images and recommend to another user by consider the image similarity grater than 0.7 and the admin can also view the recommendations, feedback, blogs of each user.
2) *Mallet:* MALLET is a Java-based package for statistical natural language processing, document classification, clustering, topic modeling, information extraction, and other machine learning applications to text. MALLET includes sophisticated tools for document classification: efficient routines for converting text to "features", a wide variety of algorithms (including Naïve Bayes, Maximum Entropy, and Decision Trees), and code for evaluating classifier performance using several commonly used metrics.
3) *LDA:* In natural language processing, latent Dirichlet allocation (LDA) is a generative statistical model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar. For example, if observations are words collected into documents, it posits that each document is a mixture of a small number of topics and that each word's creation is attributable to one of the document's topics. LDA is an example of a topic model and was first presented as a graphical model for topic discovery by David Blei, Andrew Ng, and Michael I. Jordan in 2003. In LDA, each document may be viewed as a mixture of various topics where each document is considered to have a set of topics that are assigned to it via LDA. This is identical to probabilistic latent semantic analysis (pLSA), except that in LDA the topic distribution is assumed to have a sparse Dirichlet prior. The sparse Dirichlet priors encode the intuition that documents cover only a small set of topics and that topics use only a small set of words frequently. In practice, this results in a better disambiguation of words and a more precise assignment of documents to topics. LDA is a generalisation of the pLSA model, which is equivalent to LDA under a uniform Dirichlet prior distribution
4) *User Module:* In user module first user registered in to the system and then login to the system and perform the operations. A registered user can only access the system when admin approve the user. After login to the system admin can find the travel route by just providing the source and destination. System recommend the route by considering both distance and minimum spanning tree based recommendation. Here use kruskal as minimum spanning tree algorithm. route generation algorithm to generate a new travel route from existing routes and user can also view the other user whose having the blog image similarity with the given user.
5) *Minimum Spanning Tree Algorithm:* A minimum spanning tree (MST) or minimum weight spanning tree is a subset of the edges of a connected, edge-weighted (un)directed graph that connects all the vertices together, without any cycles and with the minimum possible total edge weight. That is, it is a spanning tree whose sum of edge weights is as small as possible. More generally, any edge-weighted undirected graph (not necessarily connected) has a minimum spanning forest, which is a union of the minimum spanning trees for its connected components. There are mainly two type MST algorithms such as kruskal and prims. Here use kruskal for route recommendation
6) *Kruskal:* Kruskal's algorithm is a minimum-spanning-tree algorithm which finds an edge of the least possible weight that connects any two trees in the forest.[1] It is a greedy algorithm in graph theory as it finds a minimum spanning tree for a connected weighted graph adding increasing cost arcs at each step.[1] This means it finds a subset of the edges that forms a tree that includes every vertex, where the total weight of all the edges in the tree is minimized. If the graph is not connected, then it finds a minimum spanning forest (a minimum spanning tree for each connected component). The user can also generate travel history and blog about their trip. And user can also view another user blogs and perform keyword searching. The keyword searching is perform based on attraction of each place. Each user can communicate to another user by sending and

©IJRASET: All Rights are Reserved

receiving messages. In user module travel route exploration algorithm is used to generate a travel route according to the point of interest of user and also use a greedy algorithm in between this to generate a route of local optimal solution from a global optimal solution.

III.RESULTS AND ANALYSIS

A. Result

In this chapter mainly describes the result and analysis of recommendations. In result phase contain the various route recommendations and in the analysis part gives the evaluation of various techniques. The experimental results of the proposed system optimal travel route discovery based topic interest and image attributes for online travel route recommendation are discussed in this section. Here using the operating system of version windows 10 and platform using is c# .net. And the database created is SQL server. Proposed system using synthetic data. Synthetic data are data which are created artificially. Synthetic data are created for obtaining specific requirements or certain criteria that is specified by the user. Synthetic data are very useful for designing system of any type because this data can be used as a simulation.

The system implemented by using mainly two modules and their sub modules. The input of the system is the source and destination of a travel route provided by the user and keyword generated by the user. All processes are applied on this input. After the processing of algorithms the system generate various route in different ways. The system also recommend a feasible route to user. Here different techniques are used to generate the output. In order to generate the route from one location to another by considering both distance and MST based route generation and use kruskal as the MST algorithms for possible route generation. Next one is the attraction search here candidate route generation algorithm is used for find the route based on attraction and also use greedy algorithm to generate a local optimal solution from a global available solution. Then the system also support category search. The system provide the facility to a user getting the information about a place by providing the category, district, and state and also take the blog about that place(if available). All recommendation output are provided in result phase and the next phase is analysis. After getting the result find the efficiency of result system perform extensive experiment on synthetically generated database and plot the result of each result as a graph.

1) Keyword Based Route Recommendation (Greedy and Attraction Based)

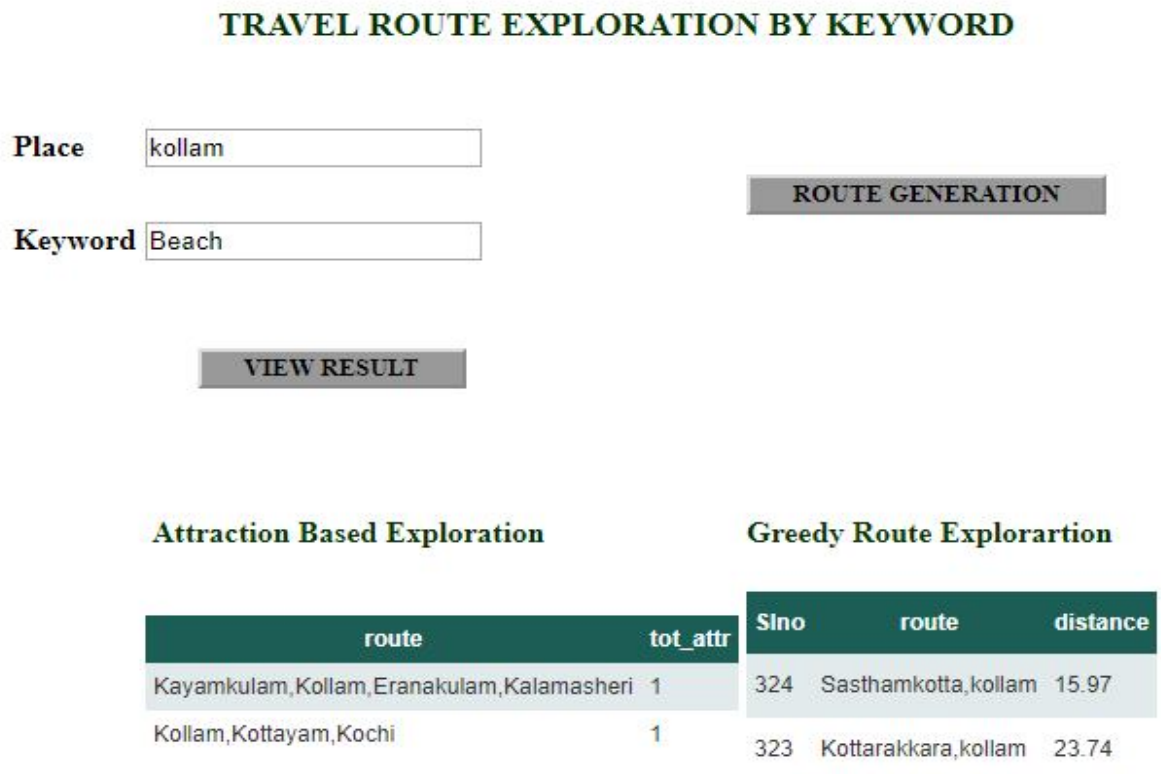


Fig 2. Recommendation based on Keyword

©IJRASET: All Rights are Reserved

2) Route Recommendation based on Distance

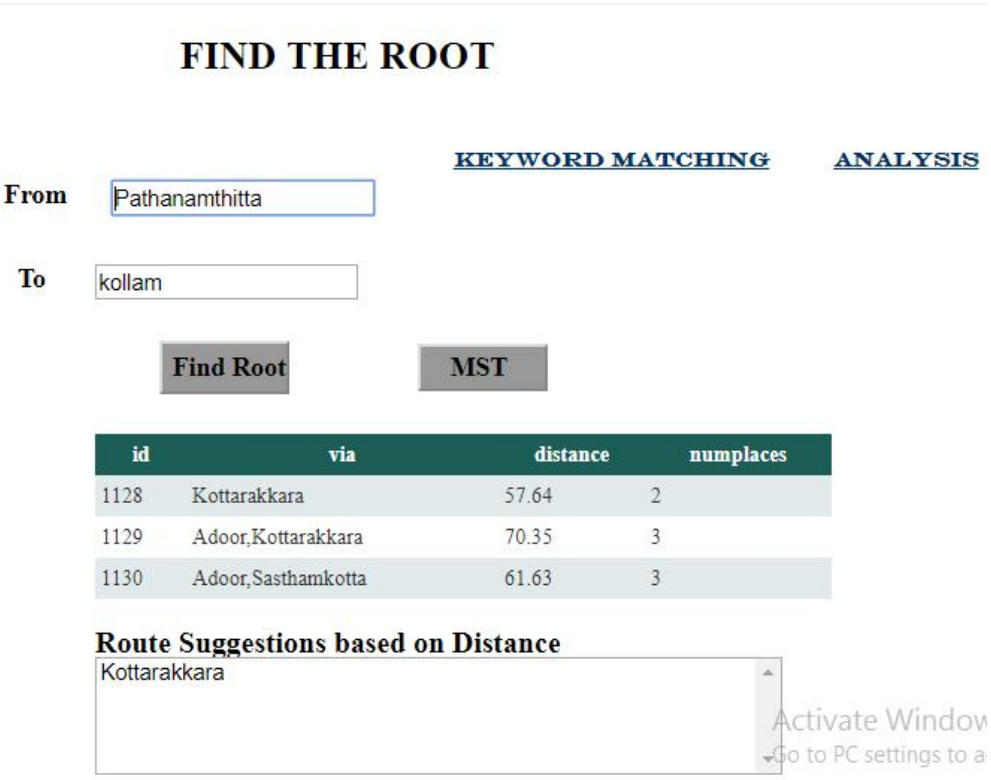


Fig 3. Route recommendation based on distance

3) Route Recommendation based on MST (Kruskal)

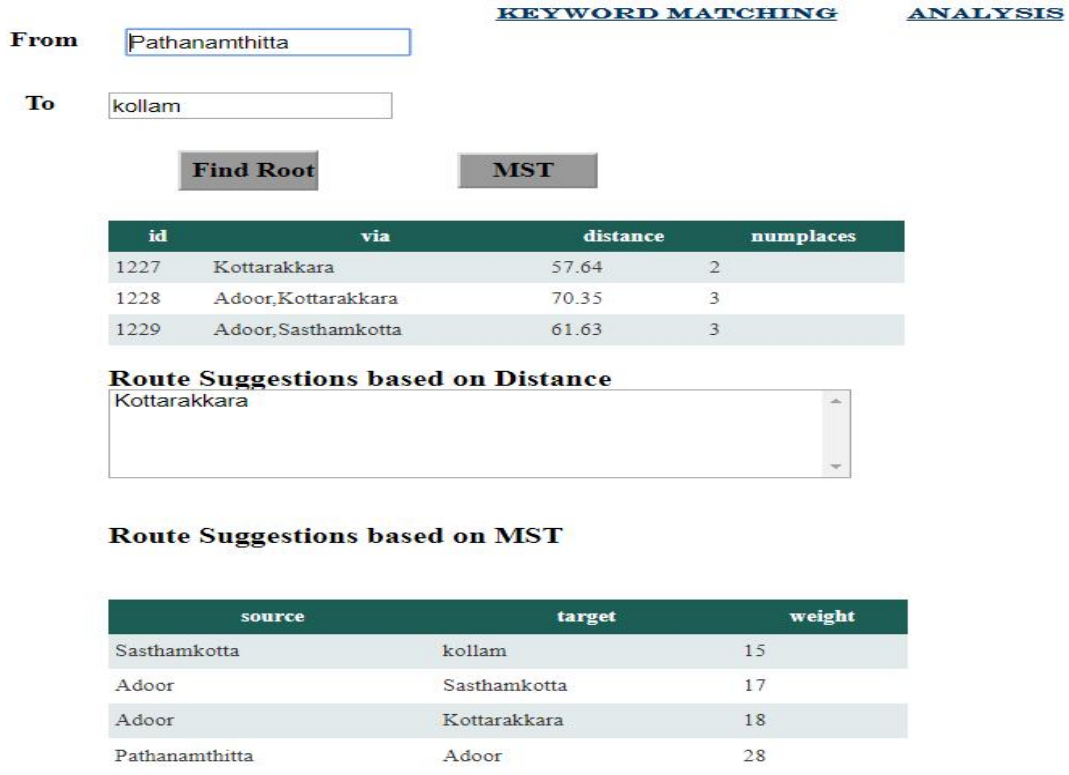


Fig 4. Route Recommendation using MST

©IJRASET: All Rights are Reserved

4) User Recommendation Based on Blog Image Similarity

Email id: aiswaryas@gmail.com

slno	suser	duser	similarity
2423	aiswaryas@gmail.com	jubin16@gmail.com	0.78828
2424	aiswaryas@gmail.com	kannan123@gmail.com	0.78828

ANALYSIS

Fig 5. Image Similarity based User Recommendation

B. Analysis

This system shows accurate travel route recommendation to users based on their preferences. It also shows analysis based on algorithms that are used in the system. Analysis also show the accuracy of getting the route with the help of different techniques. Here use minimum spanning tree, greedy algorithm, travel route exploration algorithms are used. Efficiency of the system is analysed with the help of synthetic dataset.

1) *Distance based Search and MST based Search:* This is the analysis of route recommendation by different methods such as minimum spanning tree and distance. And the graph shows the accuracy of routegeneration with the help of these methods.

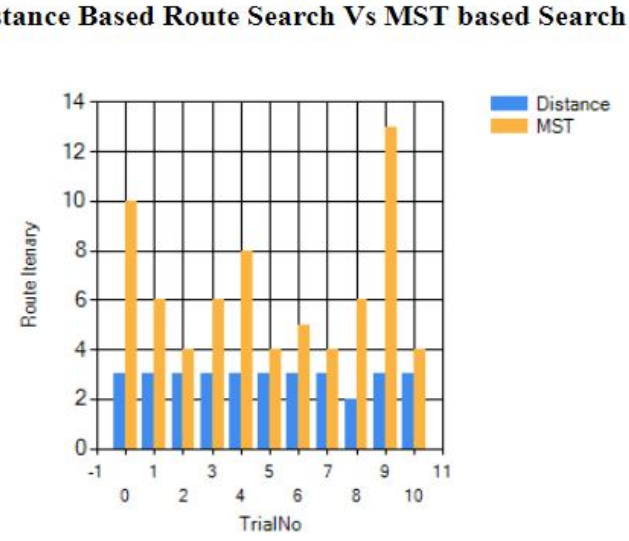


Fig 7. Distance v/s MST based Recommendation

2) *Attraction Search:* Here the analysis is taken by using the output of attraction search and greedy route generation. The attraction based route generation is done by the help of using travel route exploration algorithm. Then the local optimal solution is obtained from global optimal solution with the help of greedy algorithms. And the graph shows the efficiency of these two algorithm by analysing the output generation.

©IJRASET: All Rights are Reserved

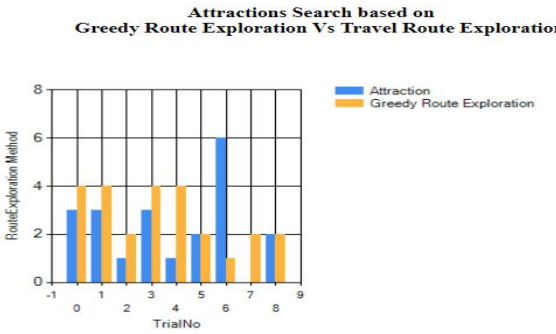


Fig 8. Greedy v/s Travel Route Exploration Recommendation

3) *User Recommendation:* In user recommendation the output is based on the number of user recommendation to another user based on image similarity. The blog image similarity is obtained by using an image comparison tool and then set a threshold 0.7 and recommend user to another when the image similarity value grater than 0.7. And the analysis shows the recommendation result of number of user recommend from a total number of users.

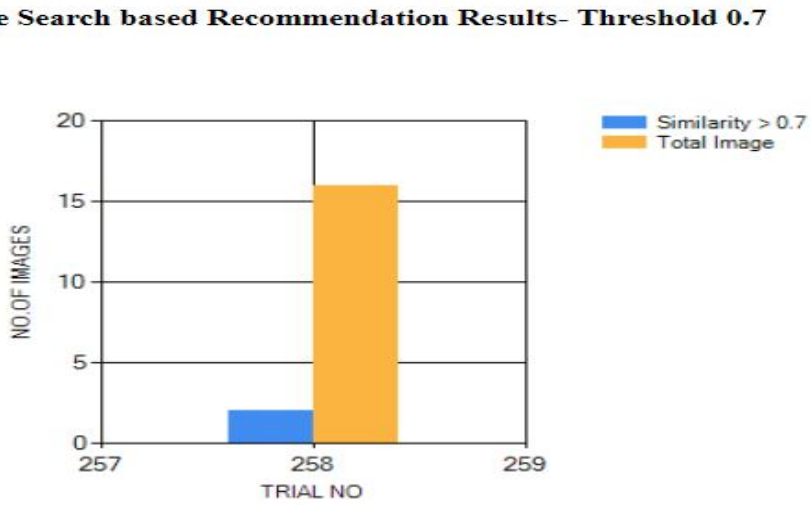


Fig 9. Image Similarity Analysis

IV. CONCLUSIONS AND FUTURE WORK

The proposed system is a travel route recommendation system, which recommend a travel route according to the user's requirement. The route generation is based on many criteria one is the distance and another is the MST algorithm here use kruskal as MST. It is very effective method for finding possible cost effective routes. The system also recommend users to another user based on image similarity. Attraction based route recommendation is also available. To evaluate the effectiveness and efficiency of the proposed algorithms, system have conducted extensive experiments location-based datasets, and the experiment results show that proposed methods do indeed demonstrate good performance compared to another works. In future use time efficient algorithm for better result and also use image as input for route recommendation.

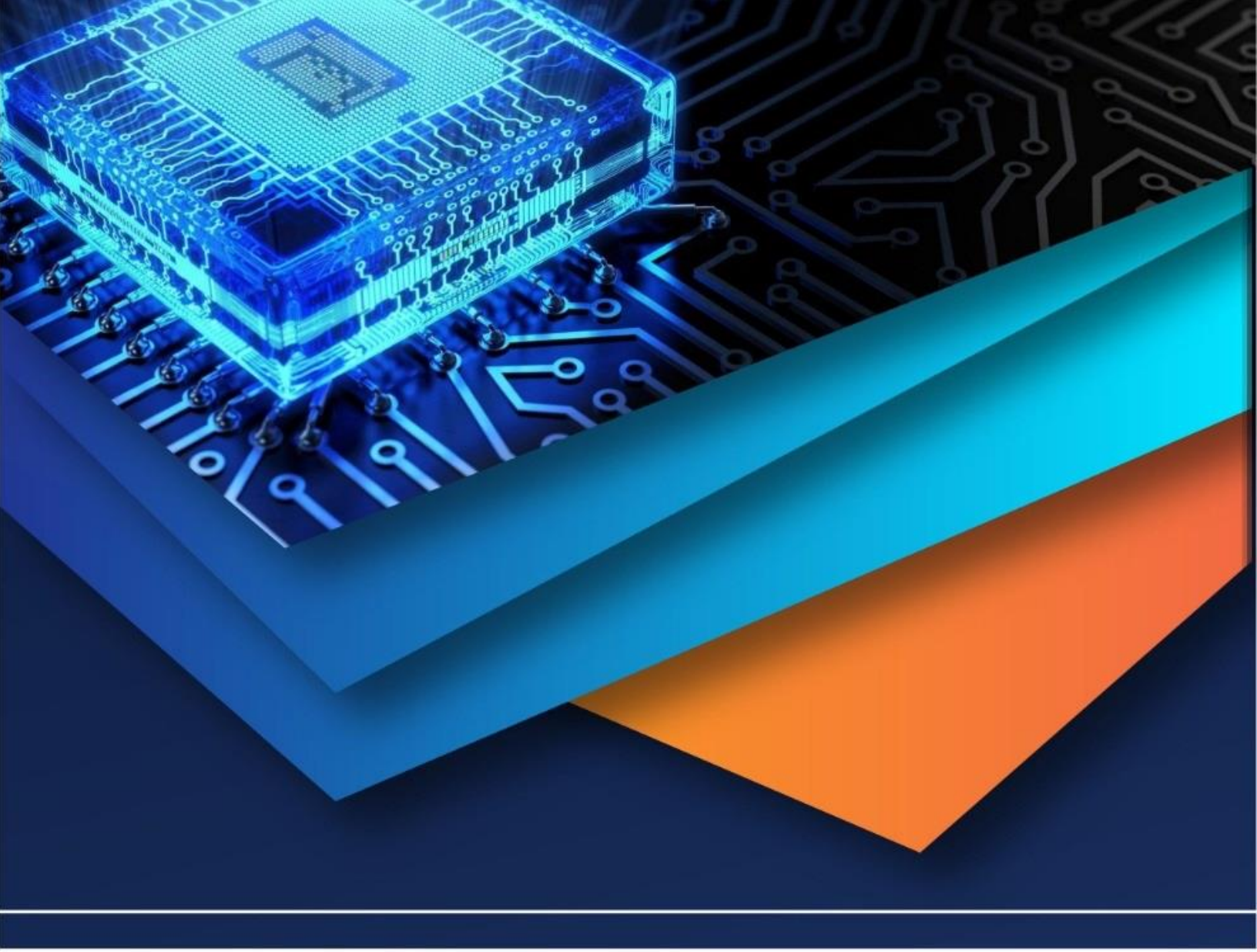
REFERENCES

[1] Yu-Ting Wen, Jinyoung Yeo, Wen-Chih Peng, Member, IEEE, and Seung-Won Hwang, "Efficient Keyword-Aware Representative Travel Route Recommendation," IEEE transactions on knowledge and data engineering, vol. 29, no. 8, august 2017 1639
[2] H.-P. Hsieh and C.-T. Li, "Mining and planning time-aware routes from check-in data," in Proc. 23rd ACM Int. Conf. Conf. Inf. Knowl. Manage. 2014, pp. 481–490.
[3] W. T. Hsu, Y. T. Wen, L. Y. Wei, and W. C. Peng, "Skyline travel routes: Exploring skyline for trip planning," in Proc. IEEE 15th Int. Conf. Mobile Data Manage. 2014, pp. 31–36
[4] Yuan, G. Cong, and A. Sun, "Graph-based point-of-interest recommendation with geographical and temporal influences," in Proc. 23rd ACM Int. Conf. Conf. Inf. Knowl. Manage. 2014, pp. 659–668

©IJRASET: All Rights are Reserved

[5] Kurashima, T. Iwata, G. Irie, and K. Fujimura, "Travel route recommendation using geotags in photo sharing sites," in Proc. 19th ACM Int. Conf. Inf. Knowl. Manage. 2010, pp. 579–588
[6] Y. Ge, H. Xiong, A. Tuzhilin, K. Xiao, M. Gruteser, and M. Pazzani, "An energy-efficient mobile recommender system," in Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2010, pp. 899–908
[7] X. Cao, L. Chen, G. Cong, and X. Xiao, "Keyword-aware optimal route search," Proc. VLDB Endowment, vol. 5, no. 11, pp. 1136–1147, 201
[8] H. Yin, X. Zhou, Y. Shao, H. Wang, and S. Sadiq, "Joint modelling of user check-in behaviors for point-of-interest recommendation," in Proc. 24th ACM Int. Conf. Inf. Knowl. Manage. 2015, pp. 1631–1640
[9] Y.-T. Wen, K.-J. Cho, W.-C. Peng, J. Yeo, and S.-W. Hwang, "KSTR: Keyword-aware skyline travel route recommendation," in Proc. IEEE Int. Conf. Data Mining, 2015, pp. 449–458.
[10] Z. Yin, L. Cao, J. Han, J. Luo, and T. Huang, "Diversified trajectory pattern ranking in Geo-tagged social media," in Proc. SIAM Int. Conf. Data Mining, 2011, pp. 980–99
[11] X. Lu, C. Wang, J.-M. Yang, Y. Pang, and L. Zhang, "Photo2trip: Generating travel routes from Geo-tagged photos for trip planning," in Proc. 18th ACM Int. Conf. Multimedia, 2010, pp. 143–152.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)