PROFICIENT ROUTE MINING APPROACH FOR TRAVEL ROUTE SEQUENCE PACKAGE RECOMMENDED SYSTEM WITH VECTOR COMPUTATION

Author : Ms.S. Hemalatha M.C.A.,M.Phil.,

Co-Author : K. Sathya M.C.A.,

ABSTRACT

The World Wide Web drags out to develop with tremendous size and multifaceted nature alongside of numerous applications. For the clients of these application progresses toward becoming tedious and hard to acquire the data they are searching so for. Such sort of user interfaces could assist them to discover the data understanding with their interests by tweaking these applications. This paper provides the suggestion of travel sequence route packages based on the preferences of user's point of interest which is associated with geographical location and tags using route mining algorithm. And after that the non illustrious and natural travel course suggestions are recommended by authentic travel records of comparable clients in the group.

In order to find the user's interests, their queries and tagging information is integrated with this approach. In this methodology, a similarity vector space model contains the information about travel sequence packages and user desired route packages are represented in such a way that the packages are suggested to the users based on the similarity. The scope of this study is to achieving the scalability and high accuracy in suggesting route travel sequence packages. Suggesting the Point Of Interest (POI) route travel package sequence with season and timing is focused on this work.

INTRODUCTION

1.1 Overview of Data Mining

Data Mining is the process of extracting information from large amount of data sets. On the other hand, it is the procedure of mining knowledge from data. The knowledge or information which is dig from the data warehouses can be used for many of the real world business applications. The data mining programs is used to analyze the patterns and associations based on the user requested data. Considering an illustration, it can be used to create labeled information classes. To demonstrate, visualize a business organization wants to use data mining to determine when they prefer to give offers and specials to the users with the certain products. The information which has collected and creates classes based on what they ordered the products and when they seen. [Reference:https://en.wikipedia.org/wiki/Data_mining] The process of data mining can be classified into the following steps. The initial step is collecting the data and loads it into the data warehouses. After that manage and the stored data could be either on domestic servers or the remote server may the cloud environment. The professionals of information technology, organization's management teams and business analysts access the data and they can organize the same based on their decidedness. Then, the stored data is sorted using the specific application software based on the results of the user required data. Finally the data can be shared and presented to the end users with various rich formats like tables or graphs which are readily available for their usage.Data mining process is entirely depends on collection of valuable data from the data warehouses along with computer process [17]. And also the data mining process is used by corporations to go round of the original data into constructing the functional and serviceable data. The business corporate uses the mining software for prototype in large amount of data to learn more about the details of development of marketing strategies and their customers to increase their sales and decrease costs.

1.2 Preparation of Data

Data cleansing and preparation is time and again abandoned but imperative steps in the process of mining the data. It is particularly applicable to the

238

typical applications collected the large data sets using automatic methods and submitted the same as the input for the analytical process in data mining. While analyzing these datasets leads many problems and it generates misleading results if the data is not carefully prepared in the data mining prediction process. Preprocessing of data in the analysis has important aspects or steps. The truthfulness of the method is depends on the quality of the data is obtained or extracted [16].
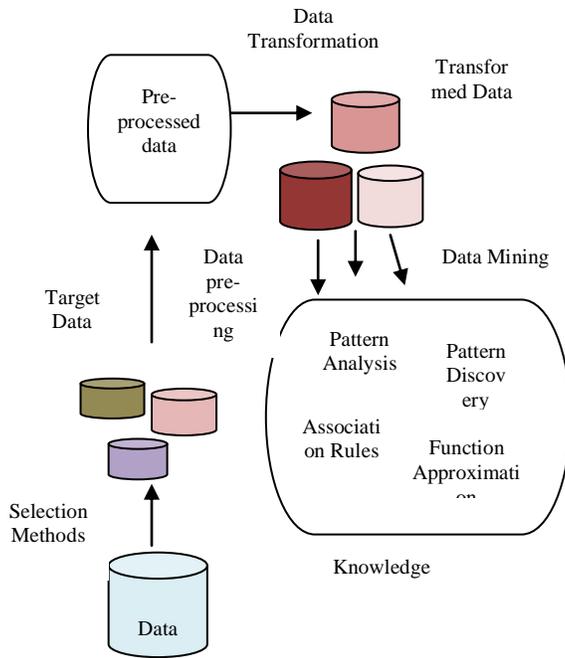

Fig 1.1 Typical Data Mining Process

The general data preprocessing step involves with the data cleansing, transformations, missing values identification and organization of data. Reduction of data in the framework of data mining is generally employed to the applications where the goal is to aggregate or amalgamate the information contained in large datasets into manageable information. It includes aggregation and simple tabulation for analyzing the principal components and clustering [15]. How many data points are covered is considered with the growth of structured and unstructured data and the speed of the data is collected and created and collected in big data from divergent sources with multiple formats.

1.3 Web Mining

The web mining explores the applications of data mining techniques into the web resources and it make easy to the further development of the data mining method to consider the specific structure of data. The content of raw data for web usage mining and the expected knowledge to be derived from it pose special challenges for the data mining application developers and researchers. During the input data are taken from mostly web server logs and other web oriented data. So that the desired results is an understanding of the user's behavior in the specific domain of information search over online, shopping sites, e-learning and other internet applications. The following are the tasks of the web mining such as finding the resources, data selection and preprocessing, generalization and analysis. The task of retrieving intended web documents is called resource finding. The retrieved information or data from the web resources are then preprocessed or cleaned with the specific contents. After that pattern discovering process is involved. For this machine learning techniques are used often for the generalization process. The patterns are discovered from the separate websites and with multiple websites aswell[http://searchcrm.techtarget.com/definition/Web-mining] Finally the discovered patterns are interpreted and validated for making the information as consistent [13]. The objective of the web mining is to obtain the reliability of information according to the users and their requirements. For the recommendation, the users knowledge and to know the exact favored travel plans and used for decision making purpose as well.

1.3.1 Web Usage Mining

The web usage mining is a type of activity in the web mining which involves the discovery of user knowledge and favorites automatically from the web server or from the online database servers. The websites, in other word e-commerce sites having the huge amount of data from the users about their interest over the products, travel packages etc. Most of the information contains in those websites regarding the user ratings, interests and their profile data with the net surfing information about various topics.

1.3.2 Web structure mining

The large volumes of information are revealed in the World Wide Web documents. While

considering the links in a web document which points to other website documents that indicates the popularity of the document or concepts. Where the hyperlinks in the document come out of a web page denotes the richness of the topics which mostly covered or arranged with that document. And also this could be compared with the bibliographical citations. An intelligent method for ranking the website or web pages take the advantage of this content used by these links to find out the similar web pages which relevant. [Reference:www.web-datamining.net/usage/]

1.3.3 Web Content Mining

Extraction of useful information from the content of the web documents is the web content mining process. The web page contains the semantic information, logical layout structures. The applications of web content mining are the extraction of relations and structure of the web pages. The web content mining approach allows the users to suggest the value added services and meta-data searching by means of extracting the web document contents [18].

1.4 Recommender System with Data Mining

The recommender system which integrates the data mining techniques to make the recommendations learned from the users using attributes of user's knowledge and actions. And these applications frequently based on user profiles for development that can be incorporated with consumption history data in geographical items [20] based on the actions during particular session and or both.For example, a person can obtain suggestions by means of pay attention with people around them have to articulate with or without the computer aided systems. If many people in organizations, institutions and other places circumstances that they take pleasure in particular things or products or if someone has to agree with recommends for given book then they may treat this as recommendations or suggestions [12].The projected work focused on recommending the travel sequence route suggestion based on the user's desired locations and travel packages. To give such kind of recommendation, the anticipated work computes the similarity of the travel packages with the user interested locations and places. To compute the travel pack similarity, the similarity vector space is applied

that is cosine similarity. While computing the match, collaborative filtering, that is user provided historical routes and fascinated route plans are considered. Based on this correspondence, the suggestion will be provided.

LITERATURE REVIEW

H. Liu et al [1] stated that the fundamental problems of location based service are localization. Many researchers are currently investigating ways to use a phone-captured image for localization as it contains more scene context information than the embedded sensors. In this study, a mobile visual localization which accurately sensing the geographic scene context according to the current approach is presented. The method proposed in this study having the capable of providing a complete set of accurate parameters about the scenes which including the actual locations of both the mobile user and perhaps more importantly the captured scene along with the viewing direction.

J. Li, X. Qian et al proposed an unsupervised image Global Positioning System (GPS) location estimation approach [2] with hierarchical global feature clustering and local feature refinement. This approach consists of the following parts such as a hierarchical structure is constructed for a large scale offline social image set with Global Positioning System information.

Tao Mei et al explored the possible of location-based service [3] to deal with an advanced recommendation problem and activity plan, which is to suggest a package of sequential activities related to user context and interest. Initially the process starts with check-in record is effective in user modeling and activity recommendation is shown. Then unique transition phenomenon category in consecutive actions from a commercial check-in dataset is finding out which recommends the sequential activities. Though this study has some uncovered issues that are not exploiting higher-order category transition patterns by considering longer check-in sessions. The price range, ratings, comments are not considered for providing the suggestions.

L. Zhang et al devised a novel method towards recommending individual location history [5]. The

experiments were made by them is on the web through the personalized friend and location recommender for the geographical information systems. An individual's visits to a geospatial region in the real world were used in this methodology for their implicit ratings on that region. In terms of the user location histories, each user a group of potential friend's community is measured the similarity and provided recommendations.

X. Qian S. Jiang et al proposed an author topic model-based collaborative filtering method [8] to facilitate comprehensive points of interest recommendations for social users. The user preference topics such as cultural, cityscape or landmark are extracted from the photos with geographic constrained textual information Through this topic model instead of only from the geo-tags (GPS – Global Positioning System locations) using that approach. The Collaborative filtering is the most well-known approach, though these approaches generally affected while providing suggestions to the users.

SYSTEM METHODOLOGY

Dataset

The dataset process is intended to obtain the travel packages details for providing suggestion. The dataset contains the details about the various cities information, important and unimportant topics, interested places, nearby locations like hotels etc and customized label information of the most wanted places etc. The city details are described and submitted to the system. Along with this, particular topics or subjects in a specified city also described. All these details were kept and maintained in a database server. To create or specify the tag information of the interested, famous and non famous locations, name for the label or tag is collected and maintained with this dataset. And also the point of interest is commented with these details. For maintaining the interested places called point of interest, the important interested points or places name is submitted. Additionally the subject or topic of the place is included with this dataset. In this topic for the interested point's formation, the initial process is describing the city information. After describing the

city, there around 20 topics for the packages is constructed with this model. In this context, the topics are defined as $T = \{t_1, t_2, t_3, ..., t_i\}$ where the $t_1, t_2...t_i$ are the different interested topics and i is the total number of interested topics listed. The point of interest is then described under the topics construction. Considering an illustration, the topic which is Garden specified under in ooty city, there would be point of interests like botanical garden and rose garden. Likewise all other point of interest were associated using the similar vector space construction model. There are several user interested travel packages for each and every point of interest. In these interests, more than one topic or subject is linked or related with one point of interest.

Construction of Packages and Subjects for Suggestion
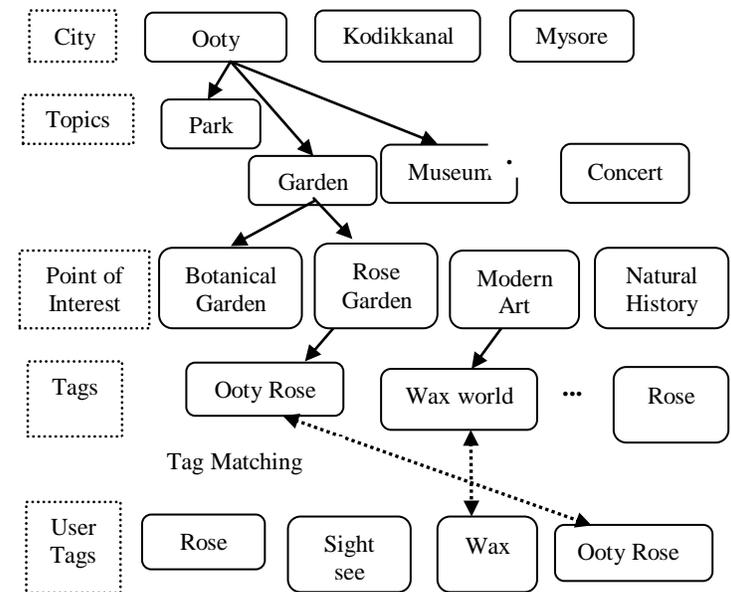
The interested topics of user's point of interest are described in similarity vector space with multiple dimensions. In this scenario, the travel sequence packages were used to extract the dataset from the different sources which may offline or online inclusive of locations. The prepared tags for the sequential travel packages are used to differentiate the topics based on the low and far above the ground occurrence while considering the user's query and tags.

Point of Interest Similarity Computation

The similarity is computed for the tags prepared for the point of interests with the user queried information and route travel packages from the dataset with user required sequence travel routes. The cosine similarity function is applied then to suggest the relevant route travel packages to the users. The cosine similarity between two vectors or two travel packages that is user required package and travel route packages is a measure that calculates the cosine of the angle among them. It is a measurement or metric of orientation and it can visualize as a comparison between travel route packages.
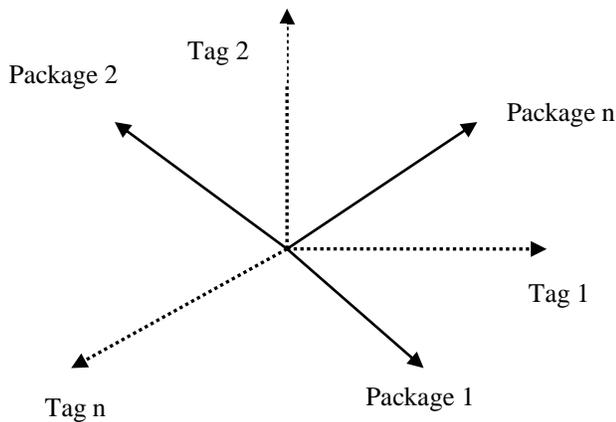
Similarity Vector Space Model

In this process to build the cosine similarity equation with the dot product for the cosθ as follows:

Package1.package2 = ||package1|| ||package2|| cosθ;
Cosθ = Package1.package2 / ||package1|| ||package2||

Computing Packages Similarity

Computing packages similarity performs the similarity calculation process for user required packages with travel packages and user query tags with point of interests (POI). For this, the sequential user package routs and travel route packages are obtained from the previous procedures. Finally, the similarity among tags or labeled marks and the user query is computed using similarity vector space model with cosθ. With this similarity score, a sequence route travel suggestion is decided. In addition to this, similarity between multiple represented lists of tags with user queries is also computed.

Pseudo-code for Sequence Route Mining Algorithm

Step 1: The point of interest of users is obtained from user query.

Step 2: According to the data of the each user's point of interest the vector structure is defined.

Step 3: Then here, historical point of interest is defined in a list and later will be defined with another list.

Step 4: Collaborative Filtering is applied based on the user's point of interest by means of counting the records from the lists.

Step 5: Sequential travel package extraction process is carried out based on the interested topics distributed($\alpha(u\ )$) and cost ($\beta(u\ )$), time ($\gamma(u\ )$)and season ($\zeta(u\ )$) for the point of interest.

Step 6: Apply the aggregate function of average of entire user point of interests to obtain the routes for the travel packages.

Step 7: The user interested topics are extracted and associated with the travel packages related to the matched point of interests distributed in the packages of topics as $\alpha(u\ ) = [\xi 1\ (u\ ), \ldots. \xi k\ (u\ ), \ldots. \xi N\ (u\ )]$.

Step 8: Resemblance function is applied then for the user's interested subject matter and routes with the travel packages which are related.

Step 9: Return the at most similarity sequential routes travel packages.

Computing Packages Similarity

Computing packages similarity performs the similarity calculation process for user required packages with travel packages and user query tags with point of interests (POI). For this, the sequential user package routs and travel route packages are obtained from the previous procedures (Section 4.2). Finally, the

242

similarity among tags or labeled marks and the user query is computed using similarity vector space model with cosθ. With this similarity score, a sequence route travel suggestion is decided. In addition to this, similarity between multiple represented lists of tags with user queries is also computed.

RESULTS AND FINDINGS

The proposed sequence travel package with relevant user approach is evaluated with the quality of similar travelling package routes generated with divergent locations and user interested points. With this performance analysis, proving that the proposed approach providing the route sequence package travels comprehensively. And also the user queried travel package route is treating as a metric for assessing this methodology and providing the similar sequence package travel. The following parameters or metrics are used for the evaluation.

 ➢ Desired location or City
 ➢ Date and Time
 ➢ Interested Points (POI)

S. No	No. of Users	Count of POI	Similarity Ratio [%]
1	5	10	37
2	10	15	68
3	15	49	87
4	20	62	91
5	30	97	92
6	50	124	95
7	100	195	97

Table 5.1: Performance Assessment of Point of Interests

Below figure describes the analysis result of point of interest (POI) with relevant package similarity ratio. The time are included in this point of interest. Along with this data, prepared labeled mark or tags for the topics in interested then extracted for this analysis from the dataset.
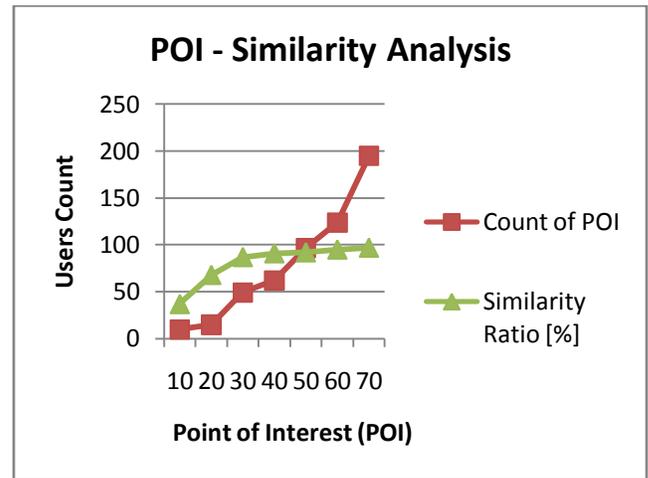

Figure 5.2: POI Similarity Analysis

The Table 5.2 describes the experimental result of the point of interest and similar travel route packages. It contains count of point of interest and matched similar travel packages count. The Point Of Interests are collected from the users and the packages are obtained from the travelogues. Using these two, similar travel Packages are returned to users.

Point of Interests	Similar Travel Packages
10	8
20	15
30	23
40	37
50	46
60	57
70	64

Table 5.2: Point of Interest and Similar Travel Packages
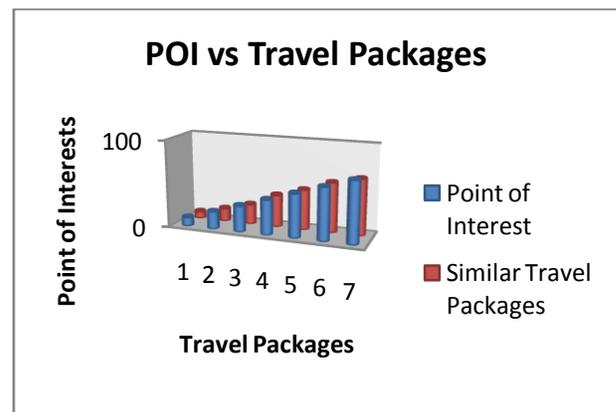

Figure 5.3: Analysis of POI vs Travel Packages

243

FINDINGS

The proposed technique gives similarity results of the travel route sequences with the user interested places. The cosine similarity vector space model computes the better match result than the preceding approaches. In this move toward more similar interested theme of the user about the places are returned which are most like and accurate. Additionally the proposed work suggests the adjacent locations along with the hotel information to the customers.

CONCLUSION AND FUTURE WORK

The anticipated proposed work helps the client to choose the succession of travel bundles to design their voyages which are connected with their purpose of intrigue. Since the approach proposed with this examination considers the many number of clients and their own enthusiasm than the personage client's purpose of intrigue. Notwithstanding that furnishes closest places with lodging, transportation and so forth. The proposed approach serves to the travel offices and the clients to set up their visits and ventures. The interface for output is designed in C#.Net windows forms with a range of controls, which make the graphical results and more satisfying. The user interested package topic's images are collectively arranged and displayed according to the user requested information. Promote course of this exploration ought to investigate with the customized and enhanced adaptation of recommendations in view of the client gave rating to the travel bundles. Another purpose of research bearing is utilizing more wellsprings of information for giving the proposal to more number of clients and input from some different visitors.

Future Work

Promote course of this exploration ought to investigate with the customized and enhanced adaptation of recommendations in view of the client gave rating to the travel bundles. Another purpose of research bearing is utilizing more wellsprings of information for giving the proposal

to more number of clients and input from some different visitors. Alongside this different pointers like instructive firms, auto/vehicle stopping, open system getting to through wifi focuses and so on could be coordinated. At long last an android and windows versatile application should be produced for this proposal framework for simple utilization of portable clients.

REFERENCES

[1] H. Liu, T. Mei, J. Luo, H. Li, and S. Li, "Finding perfect rendezvous on the go: Accurate mobile visual localization and its applications to routing," in Proc. 20th ACMInt. Conf. Multimedia, 2012, pp. 9–18.

[2] J. Li, X. Qian, Y. Y. Tang, L. Yang, and T. Mei, "GPS estimation for places of interest from social users' uploaded photos," IEEE Trans. Multimedia, vol. 15, no. 8, pp. 2058–2071, Dec. 2013.

[3] J. Sang, T. Mei, T. J. Sun, S. Li, and C. Xu, "Probabilistic sequential POIs recommendation via check-in data," in Proc. ACM SIGSPATIAL Int. Conf. Adv. Geographic Inform. Syst., 2012, pp. 402–405.

[4] Jaya Jeswani, Rajwant Gupta, Sunil Singh1, Sheena Panjabi and Ashish Tribhuvan, "Place Recommendation System", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056, Volume: 04 Issue: 03, Mar -2017.

[5] L. Zhang ,Y. Zheng, Z. Ma, X. Xie, and W. Ma, "Recommending friends and locations based on individual location history," ACM Trans. Web, vol. 5, no. 1, p. 5, 2011.

[6] Priyanka Bhalere, Sachin Walunj, Yashwant Bhaidkar, "Machine Learning Based Tourist Place Recommendation System", IJETT ISSN: 2455-0124 (Online), 2350 – 0808 (Print) , April 2016, Volume 3, Issue 1.

[7] Tielin Zhang, Yi Zeng, Hongwei Hao, "Active Recommendation of Tourist Attractions Based on Visitors Interests and Semantic Relatedness", International Conference on Active Media Technology, AMT 2014, pp 263-273.

[8] X. Qian, S. Jiang, J. Shen, Y. Fu, and T. Mei, "Author topic model based collaborative filtering for personalized POI recommendation," IEEETrans. Multimedia, vol. 17, no. 6, pp. 907–918, Jun. 2015.

[9] Yashwant Bhaidkar, Prof Sachin Walunj, Priyanka Bhalere, Pranav Bhagwat, "Tourist Place Recommendation System", ijariie, IJARIIE-ISSN(O)-2395-4396, Vol-2 Issue-3 2016.

[10] Yupeng Hu, Ruiyun Yu, Yanliang Liu and Yonghe Liu, "Application recommendation at places for mobile users", World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2016, IEEE 17th International Symposium.

[11] Rashmi Bajaj, France Telecom R&D, Samantha Lalinda Ranaweera, LSI Logic, Dharma P. Agrawal, University of Cincinnati, "GPS: Location-Tracking Technology"

[12] Subramaniyaswamy V, Vijayakumar V, Logesh R and Indragandhi V, "Intelligent travel recommendation system by mining attributes from community contributed photos" 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

[13] J. Borges and M. Levene. "A comparison of scoring metrics for predicting the next navigation step with markov model-based systems". International Journal of Information Technology and Decision Making, 9(4):547–573, 2010.

[14] J. Davidson, B. Liebald, J. Liu, P. Nandy, T. V. Vleet, U. Gargi, S. Gupta, Y. He, M. Lambert, B. Livingston, and D. Sampath. "The youtube video recommendation system". In Proc. RecSys, pages 293–296, 2010.

[15] N. D. Lane, D. Lymberopoulos, F. Zhao, and A. T. Campbell. "Hapori: context-based local search for mobile phones using community behavioral modeling and similarity". In Proc. UbiComp, pages 109–118, 2010.

First Author

Ms.S. Hemalatha M.C.A.,M.Phil.,
Assistant Professor
Department of Computer Applications,
Kongu Arts and Science College,
Erode – 638 107,
Tamilnadu,
India.

Second Author

K. Sathya M.C.A.,
Research Scholar,
Department of Computer Science,
Kongu Arts and Science College,
Erode - 638 107,
Tamilnadu,
India.