

An Effective and Customized Itinerary Planning System Using Association Rule Mining Technique with Personalized Points of Interest

M. Krishnamurthy^{1*}, D. Sudha², Y. Narasimha Rao³, R. Baskaran⁴, A. Kannan⁵

¹Department of CSE, KCG College of Technology, Chennai, India

²Department of CSE, QIS College of Engineering and Technology, Ongole, India

³Department of IT, QIS College of Engineering and Technology, Ongole, India

⁴Department of CSE, Anna University, Chennai, India

⁵Department of IST, Anna University, Chennai, India

Email: *mkrishmails@gmail.com

Received 31 March 2016; accepted 24 May 2016; published 27 May 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Mining the data from the huge collection that are present in the database and uncovering the relationships between the item set are one of the key aspects of data mining technologies. Itinerary planning system with personalization in selecting the places to the users is one of the demanding features in most of the travel plan. In this work, the system is designed in such a way to provide the customized journey plan to the users and also the effective one to the back pack travelers. Here the Points of Interests are the places to visit in each destination for the number of days chosen by the travelers. In this system, the users are allowed to specify the desired POIs to visit for the selected destination and can make their customized travel plan effectively. This proposed system is designed to choose the customized places to visit and to plan travel for K-day itineraries. The most visited itineraries are saved and updated in the database. Association rules are used to find out the frequent places visited in each destination and to provide the reputed places to the users to plan the journey. Here the Weka tool is used to evaluate the performance of the algorithm and the rules that are generated for the given travel dataset. Data set is designed by considering several attributes that can take part during travel such as source, destination, travel cost, budget, etc. Statistical analysis is done to evaluate the performance of the proposed system and the list of features that are present in the system. During the analysis part, registered users, number of logins, frequent visits, and attributes are analyzed. Thus the system can be redefined further with the

*Corresponding author.

help of this statistical analysis. It is mostly used at the organization end to evaluate their performance and improve the features. Report is generated once the user has chosen their customized places to visit and all detailed description of journey is presented to the user. Report could be saved at the user end and they can use it for the future reference. Thus the goal of the system is to provide the customized travel with personalization in choosing POIs and to find the frequent places visited with desired amenities.

Keywords

Personalized Travel, Association Rule, Statistical Analysis, Report Generation

1. Introduction

Itinerary planning is one of the most needs by all the backpack travelers and customers. Thus most of the customers prefer to plan their travel in online systems and need an effective plan. Budgetary travel is not widely suggested by the system with several places to visit. To overcome these disadvantages, this new implementation is proposed in choosing customized places to visit during the journey. Here the POIs are nothing but the places to visit in the journey by the travelers. Personalization in choosing the POIs is allowed to the travelers.

In this work, the users can plan for the K-day itinerary plan and the business visitors could choose their personalized budgetary travel. Inconsistent data are removed and stored in the database by the admin. Here the data mining technique such as association rule algorithm is implemented to uncover the relationship between the POIs and to suggest the frequent places to the customers. Association rule is widely used to find out the relationship between the items and to identify the customer behavior in purchasing the items.

Here the association rule is used to identify the user travelling patterns and the interests of the users to visit the places. Waikato Environment for Knowledge Analysis *i.e.*, Weka tool is used to generate the rules for the designed travel dataset. Statistical report analysis is done to evaluate the organization profit and loss and the reach of the system to the users. Report is generated to the customers with detailed information of the travel plan for their future reference and use. Thus with this new implementation a personalized planning of the journey could be done and satisfaction of the customers can be achieved.

In most of the existing system, the user can choose only the mentioned or availed places to visit for their trip, thus making the customers choose the option which is availed. Another disadvantage in the existing system is that the user cannot specify their budget, and only the packages based on the places will be listed to the users; thus, the users cannot mention their budget for the trip. There are so many reputed places available for the tour but still all the POIs are not enlisted to the users. Only a few number of POIs listed and most of the popular POIs are uncovered in the trip plan. Next, the distance between POIs is not summarized to the users, since the POIs may be from different countries. By considering these factors, this paper is proposed by merging the algorithms in data mining.

In this work, the user is enlisted with several POIs to visit for their trip; they can choose their customized POIs from the several places. Users can mention their budget, thus it will help the customers to choose their budgetary trip. Here the association rules are generated to know the tentative or interested places that can be visited by the users with several attribute factors involved during the trip plan. Summarization of the trip plan will be given to the user once the report is generated. The system is efficient to the user since it is designed in such a way to make the personalized travel. It is designed with the consideration of giving the customized trip to all the people by considering the cost factor. By considering most of the factors in the existing system, this system is proposed for the customized trip plan.

2. Overview

2.1. Problem Statement

This section describes about the problem definition of the customized travel planning system. POIs provided to the users by the online travel agencies are not sufficient and effective one to the customers. The motivation of

the project is to provide the personalization to the customers. There are several popular places that can be present in a country. Based on the reviews given by the several travelers given the popular POIs are uncovered and stored in the database for the effective travel plan.

Association rules are generated with random support and confidence to identify the frequent places and the amenities needed by the customers and backpack travelers. Rules are generated with the help of weka tool to save the time and to measure the performance of the system. All the places are considered and estimated with the distance to keep the nearby POIs to visit on the same day to make the effective travel and to see the several POIs with the budgetary travel. Statistical and report analysis are done to measure the profit and loss of the organization.

2.2. System Architecture

In this proposed system, the users are allowed to choose their interested POIs to visit in the travel and designed to be effective to the backpack travelers. Most travel agencies do not provide the personalization in choosing the places to visit for the journey. Only the few sub places for a particular country is suggested to the users. Thus in most of the existing systems users cannot choose their interested places to visit.

Figure 1 shows the architecture of a personalized journey planning system and explains the overall features and functionalities of the proposed system. Here the web interface is used to interact in to the system. There are three active roles present in then system. They are customer, administrator and database. Users can submit their needs and queries in the system and plan for the journey. Likewise the admin also interacts with the system through the web. The architecture is designed in such a way to uncover the relationship between the popular POIs and to allow the users to plan for k-day itinerary plan and to provide a personalized trip plan.

Users are the clients who plan for the journey, and the admin controls the overall system functionalities. Admin takes a key role and stores all the travel related information in to the system. Admin can add and delete the POIs in the database. To make a personalized travel plan, the user has to register in the system. In the registration, all the necessary details about the traveler have to be filled and stored in the database. Once the user furnished the information the data will be stored in the database.

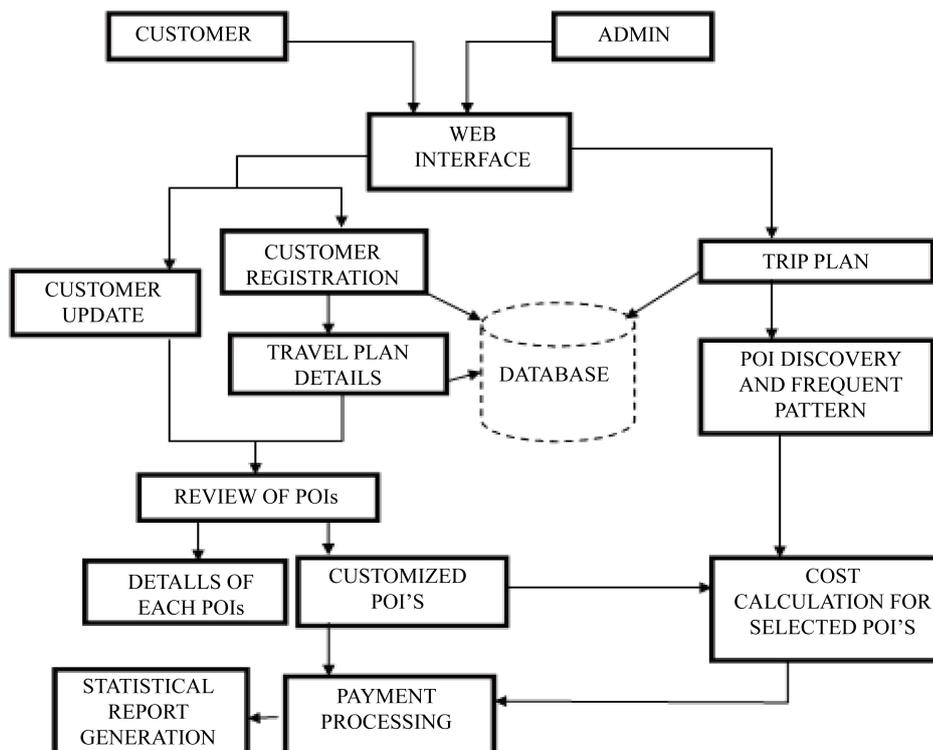


Figure 1. Architecture of personalized points of interest.

Customers update can also be done. Users are provided with the login and they can plan for the travel using the given login. Users should give the travel plan details such as the destination, budget of the travel and date, etc. Based on the cost, the system suggested the POIs and the user can select and review the places to visit. The details of each place are given to the users for their effective use of journey. After choosing their customized POIs they can go for the payment processing. At the admin end, the admin stores all the travel related information in to the system.

Frequent places and suggestion are provided to the user by the system. The admin stores all the popular POIs for each destination in the database by reviewing the several places. All the POIs are added with the necessary information and cost of each place. Thus with these detailed information user can review the places and can plan their travel.

2.3. Related Work

Pre-trip plan with dynamic information [1] are available to the users with road navigation system. Predefined information with huge collections that are stored in the storage systems and upon the request all the information is retrieved in online to the users. Travel time are estimated and calculated accurately for the online retrieval of information. Offline storage of path and online retrieval of candidate path are the features of the system in the customized travel plan system.

Traveler's decision making problem [2] widely reflects in all the itinerary planning system and resolves the problem to the users in visiting the intermediate stop during the entire journey. Personal travel companion with multimodal travelers [3] supported with existing information. But even if there was a slight change in the information it is reflected with drastic change in the real time information of the travel planning system. Optimal route planning with underlying geometric structure [4] by processing the information to answer the shortest path queries in a short time. GPS trajectories and A* algorithm used to construct and locate the travel to the users. All the domain system cannot be accommodated by this system. Hence it leads to a limitation in using the system by the users.

Travel recommendation presented to the users by mining the people attributes [5] by grasping the information that are avail in the community forums. Thus with the help of predefined information that are present already the packages are presented to the users. But it was difficult to mine all the attributed of different people, because each people may have different about the same place even though if the people attributes are relevant to each other.

A cocktail approach [6] proposed to online travel information to the users for personalized travel recommendations. Travel packages represented with different topic distributions and travel group is marked with the relationships that are present in the tourist package. Travel group is identified based on the characteristics and likings of each user that are travelled already to a place. But the system is failed to provide the preference that the user likes to have in the journey, thus the user could not make his/her personalized travel plan. Predefined itineraries are always provided to the users in most of the existing systems. In some of the itinerary planning systems, users are allowed to choose the customized itineraries but customers are limited in choosing the places to visit.

This section discusses about the features that are present in the existing systems and their functionalities. ENOSIS system [7] designed to support the travel decision to the users and presents with real time information throughout the travel plan.

Global sight map and local sight map [8] with various dimensional projections provide the various angle views of each place that are present in the journey. But the time was not effectively considered to visit the sub places that are present in the journey. ENCLUS clustering algorithm is designed to provide the personalized travel planning system. Visualization technique was quite familiar to the users and in understandable form. Single day itineraries are integrated to provide travel for multiday itineraries to plan travel for N number of days [9] and useful for the backpack travelers. Preprocessing stage is used to generate single day itineraries with the help of map reduce algorithm. Online stage is used to integrate the single day itineraries to support multiday itineraries using greedy approximation algorithm. But it takes more time to sort and group the places to visit in the travel system.

Frequent item sets to mine data from the temporal database [10] using bit vector mining algorithm. It was designed to uncover the relationships between the items in temporal database. Improved cluster based mining algorithm [11] to find the relationships between item sets using bit vector mining algorithm. By analyzing the ex-

isting systems and their features, this proposed system is implemented in such a way to provide the customized travel plan to the users.

Destination selection model based on the link flows by integrating a specific travel plan model [12] to accomplish the travel behavior of the users. Observed link set and measurement matrix are constructed. But in this model, more trip counts were not considered hence more places are not covered for the trip plan.

3. Design of a Personalized Trip Plan

This section describes about the design of the personalized travel plan and the mining of frequent places and suggestions to the users. Since day by day the amount of data increases respect to the user preferences, the admin will face difficulty in adding all the interested POIs. This issue can be overcome in the future with automated data entry extraction tools.

3.1. Add/Delete Points of Interests

In this module the user can either confirm or remove the selected POIs for their change of plan and provide ease of use. Admin can add new emerging POIs and also can delete the places. POIs are added with updated cost and days for the journey. With the implementation of this module, the user can cancel/delete the specific POI and choose some new emerging POI as they wish. Thus the customization in choosing the POIs and the satisfaction is achieved at the user end. New POIs are added at the admin end and it will be very useful to the customers.

Add/Delete POIs is allowed for,

1. Customer
2. Administrator

At the customer end, the user can delete the POI if the distance from the previously visited POI is relatively long. Thus they can add some other Places which is relatively shorter to visit and they can be benefitted with more number of places and effective utilization of the travel plan. Several places will be available for a particular destination and user can choose it with the new emerging POI and enjoy the trip services. Likewise, the new emerging POIs are added at the customer end.

Admin will add the new emerging POIs by viewing the reviews and comments given by the customers. Thus the new places are uncovered and added by the admin in the database. If any places that are not visited or no longer interested to visit by the customers, admin can delete the places. Thus only the interested and popular places are kept in the database and avail to the users.

3.2. Frequent Visits and Suggestions

This section discusses about the frequent visits and suggestions that can be provided to the users for the personalized travel plan. Here the predefined packages will be generated for the given cost and days of travel selected by the user using Association rule. Most frequently visited places suggested to users. Popular POIs for each destination represented to the customers for their ease of use.

Dataset: Association Rule Generation

```
@relation travel
@attribute source {chennai, mumbai, delhi}
@attribute destination {montenegro, montana, barcelona}
@attribute age {adult, elder}
@attribute sex {female,male}
@attribute travel_mode {ship,flight}
@attribute accomodation {no,yes}
@attribute car {no,yes}
@attribute journey {yes,no}
@data
chennai,montenegro,adult,male,ship,yes,yes,yes
delhi,montana,elder,female,flight,yes,no,yes
mumbai,barcelona,adult,female,flight,no,yes,no
chennai,barcelona,adult,male,flight,yes,yes,no
```

delhi,montana,elder,female,ship,no,no,yes
 chennai,barcelona,elder,male,flight,no,yes,yes
 chennai,montana,elder,female,ship,no,no,yes
 delhi,montenegro,elder,female,ship,no,yes,yes
 mumbai,montana,adult,male,flight,no,yes,yes

Travel dataset holds the relation, attribute and data. Here the travel attributes such as source, destination, age, sex, travel mode, accommodation, car, and journey are some of the key attributes in the travel plan. Data holds the detailed description of each attribute and the relationships among the each attribute for the travel related information. Weka tool is used in the work to generate the set of association rule for the travel dataset. With the help of these rules frequent places with personalized information and needs can be found for each destination. Weka tool is used in the work to generate the set of association rule for the travel dataset. With the help of these rules frequent places with personalized information and needs can be found for each destination.

3.3. Statistical Analysis and Report Generation

In this module report is generated both at the user end and organization end. Statistical report is generated and analyzed by the admin. With the help of this the company can analyze the number of logins, frequent visits, number of registered users. They are represented in a graph format for the easy understand and analyze by the organization. Likewise report is generated at the customer end in PDF format. Hence they can download and keep it for future reference.

Travel report is generated to the users. With the report the customer can use it for their reference and can view their travel history and all the necessary details of their trip plan. It is used to produce a document in a format which satisfies a particular human readership. Report generation functionality is almost always present in database systems, where the source of the data is the database itself. It can also be argued that report generation is part of the purpose of a Spreadsheet. The organization's profit and loss can be analyzed with the help of this report generation.

4. Database Design

The database designed for this project for the customized travel plan and the members who can use the tourism related data. Each table contains the rows and relevant attributes.

4.1. Frequent Visits Table

In this table the user relevant details are entered and stored in the database. The attributes include country ID, country name and number of visits. All the frequent visits made by the customers stored. The frequent visits table is shown in [Table 1](#).

4.2. Login Table

In this table, the user login date, and the customer details are present. Data is entered by the customer and it is added in to the database. It contains customer ID, name and login date. Whenever the user login in to the system, the details are stored and login analytics can be done by the organization. The login table is shown in [Table 2](#).

4.3. Registered Users

Here the customer details and the registered information by the customers can be drawn from the database. The

Table 1. User analytics.

Country ID	Country Name	No. of Visits
1	Montenegro	15
2	Montana	10
3	Barcelona	17

Table 2. User login table.

Country ID	Customer Name	Login Date
1	Harini	2015-03-26
2	Yamini	2015-03-21
3	Bhargav	2015-03-26

database contains customer ID, name, place, contact number, and the registered date. With the help of this the company can analyze the number of registered users and they make the travel plan. The registered user table is shown in **Table 3**.

5. Experiment Evaluations

This section discusses about the experimental results and the proposed system performance. To analyze the time efficiency of the association based methods for the travel plan and the data has been generated and mentioned in the specified elapsed time. Both the dataset generation and the statistical analysis are generated in the aspect of processing time to compute and retrieve the data.

Different set of graphs that include pie chart, line graph, bar graph are used to represent the number of frequently visited places and their comparison, number of logins per day, and number of registered users in each day to represent the statistical analysis for an organization. The list of attributes and their contribution are analyzed with the given database for the customized travel plan. With the help of this analysis the occurrence of each attribute can be analyzed and estimated with the weka tool.

5.1. Registered User Analytics

In this step, the number of registered users in the travel plan analyzed. The following tabular data **Table 4** shows the tabular data for the registered users in the company website and used to compare the number of users registered per day and the count is generated and measured.

The graph in **Figure 2** shows the generated graph for registered user. With the help of this graph the company can measure the number of user registered in their travel plan.

Here the number of logins in each day by the number of users is analyzed. The following tabular data represents the login per day table with number of users and number of logins and it is shown in the **Table 5**.

The line graph in **Figure 3** shows the number of login per day. The graph is represented with the number of logins in each day. The company can measure the number of logins in each day in their website.

5.2. Frequent Visit Analytics

This section explains about the frequent visits and the analytics performed by the system. The tabular representation in **Table 6** shows the tabular data for frequently visited places. It is represented with the places and the percentage level of most visited places in the travel plan.

The graph in **Figure 4** shows the frequent visits of each place in the travel plan. It is represented with the places and the percentage level of visits. Each places represented with difference colors and percentage level of most visited.

$$\begin{aligned}
 \text{Age} = \text{adult} \wedge \text{ravel_mode} = \text{flight} &\rightarrow \text{car} = \text{yes} \\
 \text{Destination} = \text{Montenegro} \wedge \text{Car} = \text{yes} &\rightarrow \text{travel_mode} = \text{ship} \\
 \text{Travel_mode} = \text{flight} \wedge \text{accommodation} = \text{no} &\rightarrow \text{car} = \text{yes} \\
 \text{Destination} = \text{Barcelona} \wedge \text{travel_mode} = \text{ship} &\rightarrow \text{car} = \text{no} \\
 \text{Car} = \text{yes} \wedge \text{age} = \text{elder} &\rightarrow \text{travel_mode} = \text{flight}
 \end{aligned} \tag{1}$$

Equation (1) shows the lists of rules generated by the weka tool for the given travel dataset with relation, attribute and data. With the help of these rules frequent needs of the user can be seen and the organization can improve their functionalities.

Table 3. Registered user table.

User ID	User Name	Place	Contact Number	Registered Dates
1	Yusuf	Chennai	9789541296	28-12/2014
2	Yamini	Chennai	9500254789	29/12/2014
3	Gaurav	Chennai	9600154789	29/12/2014

Table 4. User analytics.

Country ID	Customer Name	Login Date
1	Harini	2015-03-26
2	Yamini	2015-03-21
3	Bhargav	2015-03-26

Table 5. Logins per day.

Number of Users	Date of Login
1	2015-03-02
1	2015-03-21
1	2015-03-27
2	2015-03-20
3	2015-03-26

Table 6. Frequent visits table.

Points of Interest	Percentage
Montana	36%
Barcelona	24%
Montenegro	40%

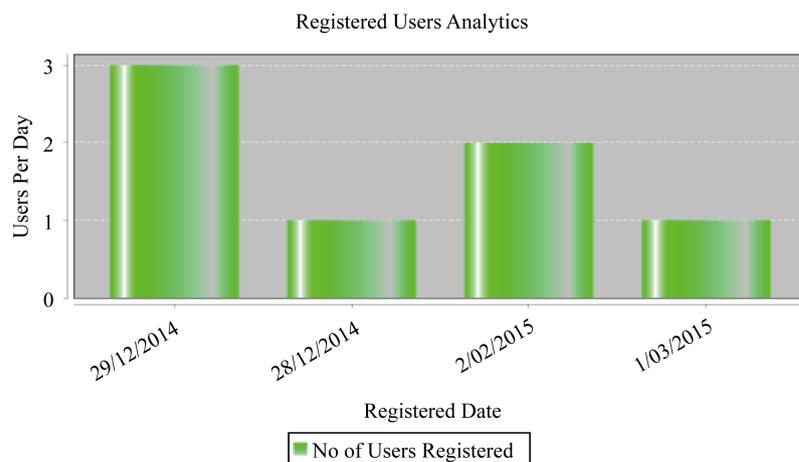


Figure 2. Statistical user analytics.

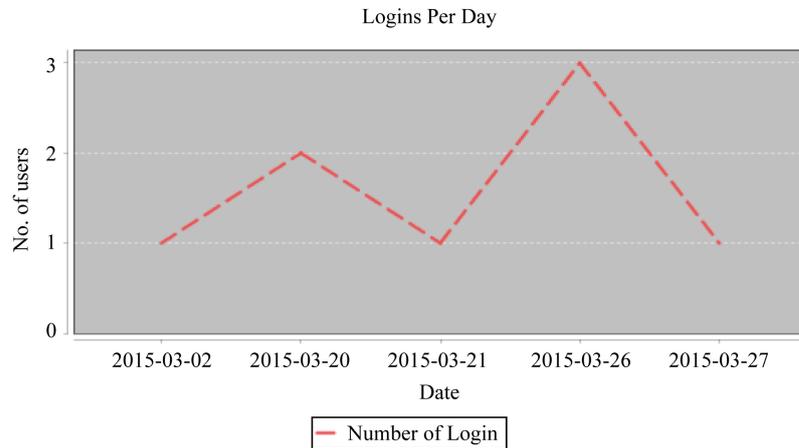


Figure 3. Statistical login analytics.

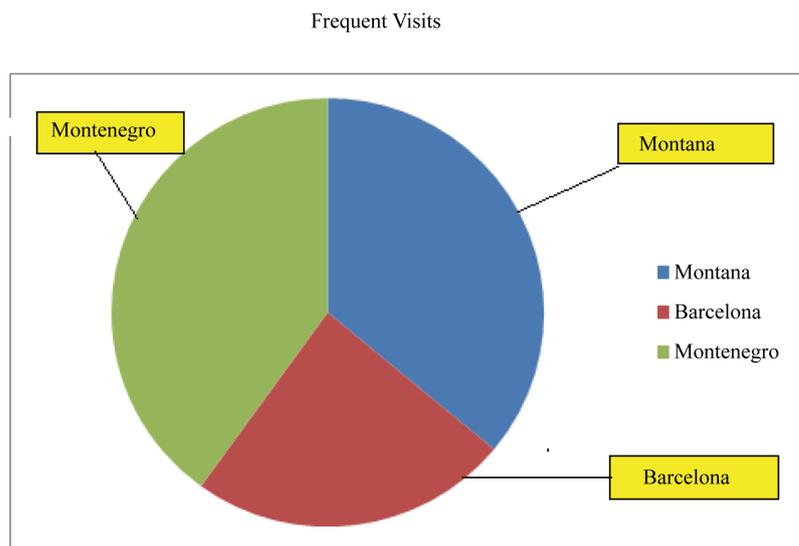


Figure 4. Frequent visits generation.

5.3. Attribute Analysis

This section explains about the analysis of attribute present in the travel dataset. **Table 7** shows the tabular representation attribute table.

Figure 5 shows the attribute collections for the travel data with different attributes to make customized travel. The attributes are source, destination, age, sex, travel mode, accommodation, car and the class journey. Here the dark blue represents the positive and the red color represents the negative option. The color differentiation helps to know the importance of each attribute.

Figure 6 shows the visualization of data in the cluster form. Each cluster is a group. A group contains a set of similar data. Here each group contains the data about the user’s fields such as POIs, budget, date of travel, list of countries, etc., such that each cluster holds the similar data about the user’s fields.

Figure 7 shows the outliers detection, which represents the data that resides out of the boundary or from the desired function. This can be related with list of POIs that resides out from the each destination.

The above table shows the three main attributes and its features involved in the personalized travel. The attribute represents the Points of Interest which allows the user to choose their desired places to visit, budget which allows the user to specify their cost of journey and travel mode which allows the user to choose their mode of objects that can be availed during travel.

Table 7. Attribute table.

Attribute	Yes	No
Source	6	4
Destination	9	4
Age	6	5
Travel Mode	6	7

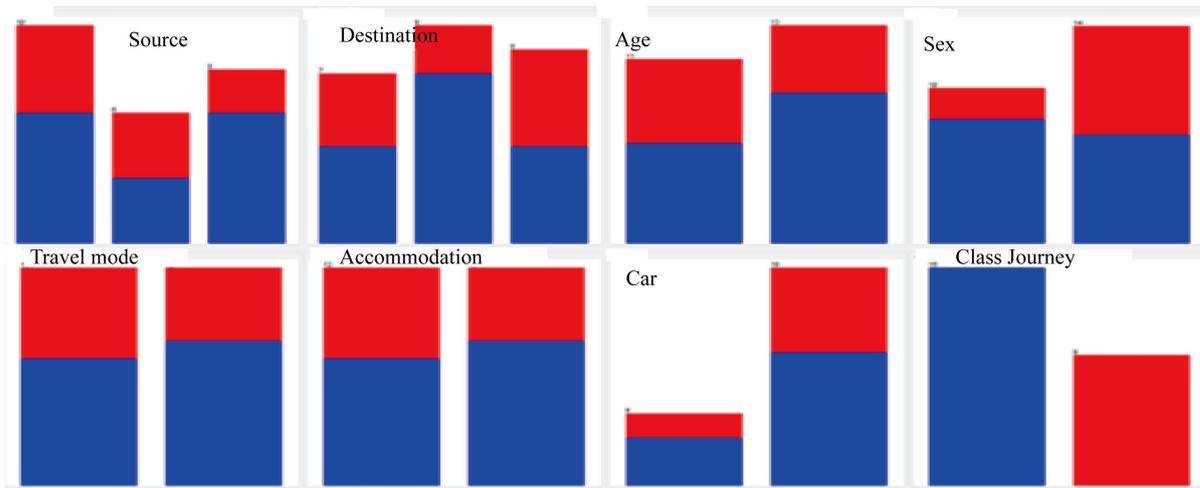


Figure 5. Attributes representation.

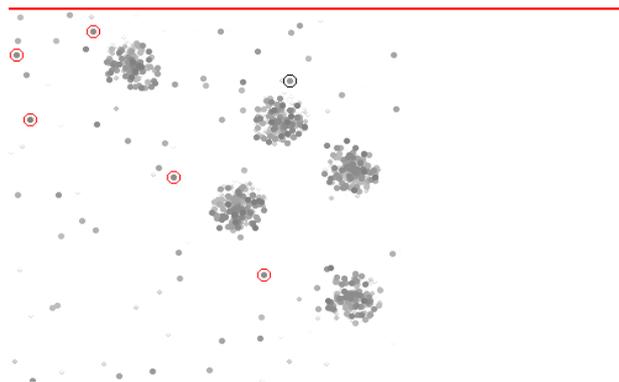


Figure 6. Visualization of data.

6. Conclusions

In this work data mining algorithm is implemented to find the frequent places. To evaluate the frequent places, association rules are generated. Dataset is created and association rules are automatically generated with the help of weka tool. Weka is widely used to perform the data mining tasks of data mining algorithm and machine learning algorithms. Weka tool holds several algorithms such as classification, clustering, preprocessing, association rules, visualization techniques, etc. With this new formulation users can choose the customized places to visit for their journey. Dataset is designed with relation, attributes and the data that can be used in the travel.

Association rule gives the suggestions to the travelers. This new proposed system is designed to meet the needs of the customers and back pack travelers. At the organization end to evaluate the use of the system, statistical analysis is done. Statistical analysis includes the number of registered users, number of logins per day and the frequent visits of each destination. New frequent POIs are stored and the POIs which are not visited by the

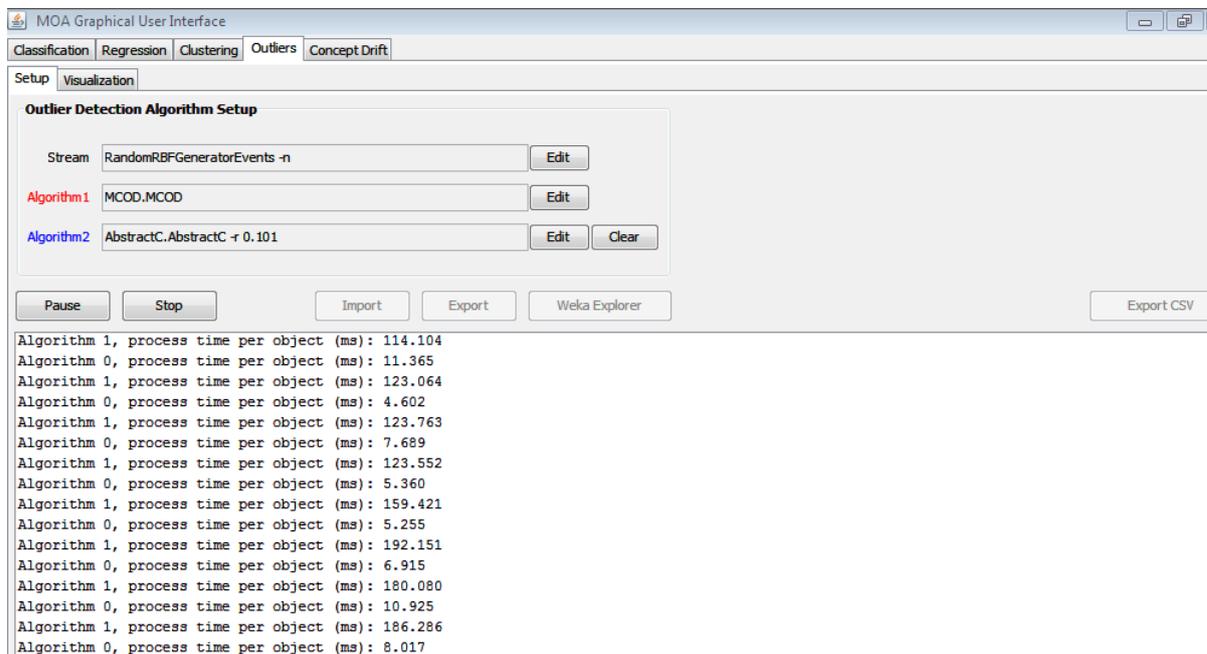


Figure 7. Outliers detection.

customers are also deleted by the admin. Report is generated at the customer end to list all the necessary information and the POIs that the user selected for the journey. User can save it and use for the future reference. By this proposed implementation, customized travel plan could be achieved and the satisfaction can be seen at the customer end. In the future enhancement, this system can be proposed with big data analytics to manage the several POIs and data present in the system.

References

- [1] Chen, Y.Y., Bell, M.G.H. and Bogenberger, K. (2007) Reliable Pretrip Multipath Planning and Dynamic Adaptation for a Centralized Road Navigation System. *IEEE Transactions on Intelligent Transportation Systems*, **8**, 14-20. <http://dx.doi.org/10.1109/TITS.2006.889437>
- [2] Zografos, K.G. and Androusoyopoulos, K.N. (2008) Algorithms for Itinerary Planning in Multimodal Transportation Networks. *IEEE Transactions on Intelligent Transportation Systems*, **9**, 175-184. <http://dx.doi.org/10.1109/TITS.2008.915650>
- [3] Rehrl, K., Bruntsch, S. and Mentz, H.-J. (2007) Assisting Multimodal Travelers: Design and Prototypical Implementation of a Personal Travel Companion. *IEEE Transactions on Intelligent Transportation Systems*, **8**, 31-42. <http://dx.doi.org/10.1109/TITS.2006.890077>
- [4] Edelkamp, S., Jabbar, S. and Willhalm, T. (2005) Geometric Travel Planning. *IEEE Transactions on Intelligent Transportation Systems*, **6**, 5-16. <http://dx.doi.org/10.1109/TITS.2004.838182>
- [5] Chen, Y.-Y., Cheng, A.-J. and Hsu, W.H. (2013) Travel Recommendation by Mining People Attributes and Travel Group Types From Community-Contributed Photos. *IEEE Transactions on Multimedia*, **15**, 1283-1295. <http://dx.doi.org/10.1109/TMM.2013.2265077>
- [6] Liu, Q., Chen, E.H., Xiong, H., Ge, Y., Li, Z.M. and Wu, X. (2014) A Cocktail Approach for Travel Package Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, **26**, 278-292. <http://dx.doi.org/10.1109/TKDE.2012.233>
- [7] Zografos, K.G., Androusoyopoulos, K.N. and Spitsadakis, V. (2009) Design and Assessment of an Online Passenger Information System for Integrated Multimodal Trip Planning. *IEEE Transactions on Intelligent Transportation Systems*, **10**, 311-323. <http://dx.doi.org/10.1109/TITS.2009.2020198>
- [8] Nam, J.E. and Mueller, K. (2013) TripAdvisor N-D: A Tourism-Inspired High-Dimensional Space Exploration Framework with Overview and Detail. *IEEE Transactions on Visualization and Computer Graphics*, **19**, 291-305. <http://dx.doi.org/10.1109/TVCG.2012.65>

- [9] Chen, G., Wu, S., Zhou, J.B. and Tung, A.K.H. (2014) Automatic Itinerary Planning for Traveling Services. *IEEE Transactions on Knowledge and Data Engineering*, **26**, 514-527. <http://dx.doi.org/10.1109/TKDE.2013.46>
- [10] Krishnamurthy, M., Kannan, A., Baskaran, R. and Kavitha, M. (2011) Cluster Based Bit Vector Mining Algorithm for Finding Frequent Itemsets in Temporal Databases. *Journal of Elsevier on Procedia Computer Science*, **3**, 513-523. <http://dx.doi.org/10.1016/j.procs.2010.12.086>
- [11] Krishnamurthy, M., Rajalakshmi, E., Baskaran, R. and Kannan, A. (2013) Improved Cluster Based Mining Algorithm for Frequent Itemsets Generation Using Bit Vectors. *TQ-Research Journal*.
- [12] Ye, P.J. and Wen, D. (2013) A Study of Destination Selection Model Based on Link Flows. *IEEE Transactions on Intelligent Transportation Systems*, **14**, 428-437. <http://dx.doi.org/10.1109/TITS.2012.2220135>