Tourist Spot Recommendation System Using Fuzzy Inference System

Haymontee Khan, Noel Mannan, Shahnoor Chowdhury Eshan, Md. Mustafizur Rahman,
K. M. Mehedi Hasan Sonet, Wordh Ul Hasan, Rashedur M. Rahman
Department of Electrical and Computer Engineering, North South University
Plot-15, Block-B, Bashundhara, Dhaka-1229, Bangladesh
haymontee.khan@northsouth.edu, noel.mannan@northsouth.edu, eshan.world@gmail.com, rahman.mustafizur@northsouth.edu,
sonet.hasan@northsouth.edu, wordh.u.hasan@ieee.org, rashedur.rahman@northsouth.edu

***Abstract*—A fuzzy logic based recommender system is proposed in this research that uses Fuzzy Inference System (FIS) to recommend tourist spots to travelers. The system keeps a database containing tourist locations and their respective information as well as some specific data which is used as parameters for the inference system. The system prompts user to give their desired location type(s) that they wish to visit, their budget, the number of people going on the trip, and other defined parameters which are used in the inference system. We have decided to limit the compound location type to two for simplicity of the calculation. Besides, a tourist is not normally interested in more than two location types. In addition to that, we will suggest the user some other places of different types around the selected locations which they may wish to visit during their stay. Once the input is taken, the system will fetch corresponding data from the database, fuzzify the metadata and pass it to the fuzzy recommendation engine. The recommendation engine will calculate and assign a crisp value to each location as an output. Then the list is sorted in a descending order and the top locations that satisfy most of the user's defined preferences are recommended to the user.**

Keywords-Fuzzy Logic; Fuzzy Inference; Recommendation System; Travel Agencies; Tourist Spot Searching; Tour Type.

I. INTRODUCTION

Tourism is one of the rapidly growing businesses of the world and it is contributing relentlessly in the global economy. The objective of the proposed work is to recommend places to visit within South Asia for tourists and travelers in order to assist them in planning and choosing places that are the most suitable for them and meet their needs. Nowadays, when planning a trip, regardless of the duration, most people resort to the internet. They have begun to comprehend the convenience with which they can plan a tour by properly utilizing the currently available technologies. Information about tourist spots across the globe is readily available, courtesy goes to the ever-increasing number of travelling agencies. However, travelers usually have a limited knowledge of the place they visit. They are not adequately informed about the local artistic, historical or cultural sites, the activities and entertainment options that are available to them. Neither do they fully know which other nearby places can be visited during their stay at a certain location. Travel agencies just provide them the packages, which they (agencies) prepare for the sake of their business and profit under the veil of assisting a traveler in tourist spot selection. They do not even consider the specific preferences of a traveler using which they could suggest a different spot, which would meet his/her requirements; instead they offer their predefined packages and thus fail to meet the traveler's need to a great extent. Previously other recommendation systems have been proposed and developed for recommending tourist sites, but they are limited to user preferences and user demographic profile only. Other studies have developed recommender systems by using a priority system, where a priority value is assigned to each location based on how frequently those places have been visited by other travelers. However, our proposed recommendation system will take into account the user's preferences of location types and their budget, and suggest the best-suited places to visit for their trip. The system also takes into account some attributes of the place, such as beauty, safety, facilities available to the tourists, and type (i.e. beaches, forests, historical sites) to determine the optimal solution that best suits for the user.

The rest of the paper is organized as follows. Section II presents some research findings that are related to our work. Section III describes the input variables used, Section IV explains the methodology, Section V analyses the results and Section VI draws the conclusion and discusses future implementations.

II. RELATED WORKS

The authors in [1] used a crowdsourcing based recommendation engine for locations within a city. The user may inquire about a location type within the city of Delhi, India. The user inputs their current location and their desired location type (i.e., Shopping Mall, Internet Cafe etc.) and the system gives an ordered list of such locations nearby, which the user can visit without any inconvenience.

In [2], simple fuzzy inference system based recommendation system has been designed for providing the user with the best possible choices of hotels to stay in the city of Shiraz, Iran. To facilitate the user in choosing the best hotels, the system takes different inputs from the user including preferable budget for one-night stay, facilities and preferable distance from a selected type of place. They have provided the user with the flexibility of choosing the desired distance from different centers like academic, industrial, pilgrimage, cultural, trading. To calculate overall facilities, they took into account availability of Internet, parking, restaurant etc. The authors in [3] designed the recommender system that is related to our proposed work. In this paper, a recommender system is proposed, which takes into account the user's preferences, their demographic profile and previous tours, and cross-references them with frequently visited places by others to recommend the best-suited places for the user to visit. The system maintains a database of locations and their respective information (i.e. location name, type, rating etc.) and user information (i.e. user profile, previous trips, current preferences etc.).

Authors in [6] used a hybrid approach along with K-means clustering algorithm to help tourists in determining an optimal itinerary plan. K-means algorithm was used to cluster various tourist spots, which were then used to find the optimal route among each clustered tourist spots by using greedy and 2-opt algorithm. Web contents were recommended in [7] for a particular user based on different time of the day. Tourists arrival was predicted in [8] by introducing a new approach called electromagnetism like mechanism (EM) algorithm. A Neural Network model was developed with the help of EM algorithm for learning the rule and the complete model was then named EMNN (Electromagnetism like mechanism Neural Network). The proposed model outperforms all other predicting model like multiple regressions (MR) and moving average (MV).

Three recommendation techniques like Collaborative Filtering (CF), Content Based Filtering (CB) and Demographic Filtering (DF) have been used in [9] to look at the predictive power on which the neural network model could deliver the more accurate result to suggest tourist spots. The authors noticed that nearest neighbors played a key role in this task of recommendation of spots. If the KNN value is less than 10, then the others methods seemed to produce low error values except for CB method. For the DF method, the KNN value is equal to 1, produced the best result because of the fact that similar type of respondents might belong to the same group and it generated more stable values. Even for the CF, the recommendations were better for small number of neighbors. Most importantly, the proposed hybrid Neural Network model generated the most accurate result out of all the methods irrespective of all the facts considered for other single methods. The authors in [10] designed a system that will propose stock trading decisions based on both quantitative factors, like technical indexes and qualitative factors like macro-economic or political effects. The authors in [11] focus on a type of recommendation system, dubbed by the authors as reclusive. This system as opposed to collaborative filtering approach, takes into account the similarity between objects rather than other people's preferences. The authors in [12] introduce a hybrid fuzzy recommender system for Telecom products. The authors combined user based and item based collaborative filtering techniques with fuzzy set theory to achieve an optimal recommendation engine that suggests products to users.

Researches have been conducted on this topic before, but there are several shortcomings. For instance, in earlier works, researchers had considered a limited number of factors, as opposed to our numerous variables. The analysts of previous papers considered factors limited to user demographic, where suggestions were provided based on the feedback of other users of similar demographic. Moreover, recommendations were often provided simply based on the frequency at which a place had been visited. Approaches such as these, leave much room for improvement, which is exactly what we have attempted to exploit. Our proposed work offers a multitude of variables to the user and by using those, a customized and personalized recommendation can be provided.

TripAdvisor [4] and Expedia [5] are both travel sites that offer information, such as hotel names, restaurant names, places to visit, etc. about tourist spots. They serve as information banks rather than recommender systems that provide suggestion tailored to the users' need. In both the sites, the user can enter their desired location of travel, the departure and return dates, the number of people travelling and the number of required rooms; and based on that information the site offers a list of hotels ranked according to popularity. The users can look up nearby restaurants or activities to take part in, which are all displayed based on popularity.

III. THEORY

The main objective of this research is to build a recommendation system which can provide the users the best-suited places according to their demand before they go on a travel. To understand specific requirement provided by the user, 6 different variables are used and as the output top scored locations are suggested along with other recommended places around that location. The whole system starting from taking input to processing output is described below.

A. Input Variables:

We have considered 6 input variables which act as decisive factors in getting a desirable location for a specific user under specific constraints. They are Location Type(s), Beauty, Safety, Facilities, Budget, and Tour Type. Details of these variables are provided below:

Location Type(s): Location Type is the variable that defines what kind of location the traveler wants to visit. It is used to fetch the related location data from the database. This variable is also used to calculate each location's score for the user's selected type. This variable can be a single input, e.g., Beach. User can also give multiple inputs, e.g. mountain, cave, and forest. Then, all the locations from the database, which are compatible with the user's choice are retrieved and their scores for the chosen location type are calculated. For example: assume, a user selects type forest, then all locations with type forest are fetched from the database and their scores are calculated by:

L1_Score_Type = 10 - L1_Score_As_Forest
L2_Score_Type = 10 - L2_Score_As_Forest
.
.
Ln_Score_Type = 10 - Ln_Score_As_Forest

There is also an option for the user to enter a compound type, i.e., Hills and Beaches. A location can contain both hills and beaches, but for a certain type, it is more compatible. For example, Cox's Bazar, Bangladesh is more preferable for beaches though it contains hills. Therefore, as a beach type it will score more than that type of hills. In this case, every location's scores which have location types containing both hills and beaches are added together and subtracted from 20 (two location types) to get their respective membership value. For example: assume a user selects type Hills and Beaches, then all locations with type containing Hills and Beaches are fetched from the database and their scores are calculated by:

L1_Score_Type = 20 - (L1_Score_As_Hills + L1_Score_As_Beach)
.
.
Ln_Score_Type = 20 - (Ln_Score_As_Hills + Ln_Score_As_Beach)

TABLE I. LOCATION

Location				
Name	**Country**	**LID**	**Beauty**	**Safety**
Cox's Bazar	Bangladesh	C1	7	8
Kuakata	Bangladesh	K1	5	5
Bandarban	Bangladesh	B1	8	5

TABLE II. LOCATION TYPE

Location Type		
LID	**LType**	**LRank**
C1	Beach	9
C1	Hills	6
B1	Beach	0
B1	Hills	7

From TABLE I and II, we can see that Cox's Bazar contains location type as both Hills and Beaches. So as a result, Cox's Bazar will score 15 considering both of its types. However, for Bandarban, it will score only 7 as there are not any beaches in that location. If we subtract the score according to the above-mentioned formula then Cox's Bazar will score 5 as Location type whereas Bandarban scores 13. So the lower the difference is, the higher the priority is given for that location to the user.

Beauty: It is one of the main (unique) attributes that is used as a parameter to the fuzzy inference system. Beauty simply refers to how beautiful the location is. This is a fuzzy variable with 3 levels with respect to the membership function. The levels are High Similarity, Moderate Similarity, and Low Similarity. The user is asked to enter their desired level of beauty. In our database, there is already a predefined beauty value for every location. We will just compare the beauty preference of the user with the predefined value by taking the absolute difference of that two. The lower the difference is, the higher the Similarity will be. When the related locations are fetched from the database, their respective beauty values are calculated by using TABLE I for the corresponding Beauty value of a specific location:

L1_Score = | UserPreference - L1BeautyScore |
L2_Score = | UserPreference - L2BeautyScore |
.
.
Ln_Score = | UserPreference - LnBeautyScore |

The Beauty value of Cox's Bazar is 7 but the user entered Beauty value 9 as his/her preference, so according to this above formula, Cox's Bazar will score $|9-7| = 2$ which means the location can be considered to be Similar to the user's preference level of beauty. In this case, the lower the difference is, the higher the preference will be.

Safety: The safety value is another unique attribute that is used as one of the parameters of the fuzzy inference system. Safety refers to the overall safety of the location. It takes into account the following - how safe the location is in terms of weather stability, safety based on the condition of the roads, natural disaster occurrence frequency, level of petty crimes in the area and generally how safe it is for tourists to visit the location (the general sentiment of the indigenous population towards foreigners). This is a fuzzy variable with 5 levels- Extremely Secure, Highly Secure, Moderately Secure, Highly Insecure and Extremely Insecure.

L1_Score = | UserPreference - L1SafetyScore |
L2_Score = | UserPreference - L2SafetyScore |
.
.
Ln_Score = | UserPreference - LnSafetyScore |

The user is asked to specify their required level of safety. After the related locations are fetched from the database, each safety value is calculated similarly as the beauty value is done.

Facilities: Facilities is a fuzzy variable, which is a parameter for the fuzzy inference system. This is a compound variable that contains three sub variables. These are - Hotel Availability, Transportation Availability and Recreational Activities. For each of the locations, there are predefined values for Accommodation and Transportation facilities and are kept in FACILITIES-1 table of our database (TABLE-III). However, the Recreational activity is again divided into some other activities namely Paragliding, Parasailing, Trekking etc. and those information are kept in our FACILITIES-2 table of our database (TABLE-IV). The user is asked to enter their choice of required levels of facilities. The first two are taken as regular inputs. For the third one, the user is provided with a checklist of five recreational activities. The user will provide their preferable score on a scale of 10 for the each of the facilities of TABLE-III. But, for the recreational activities, it is optional to choose any of them or all of them. The number of activities they choose will be multiplied by 10 and be added with the preference score of the Accommodation and Transportation facilities of the user. The summation is negated from the maximum possible score of a particular location to calculate the score for facilities of a location.

TABLE III. FACILITIES-1

Facilities 1		
LID	**Accommodation**	**Transportation**
C1	9	7
B1	7	5

TABLE IV. FACILITIES-2

Facilities 2		
Activity	**LID**	**Rank**
Paragliding	C1	8
Paragliding	B1	0
Trekking	C1	0
Trekking	B1	9
Parasailing	C1	7

UD = User Defined

For instance, if none of the activities are selected:

L1_Score_Facilities = | (UD_L1_Accomodation_Score + UD_L1_Transportation_Score) - (L1_Accomodation_Score + L1_Transportation_Score) |

.

.

Ln_Score_Facilities = | (UD_Ln_Accomodation_Score + UD_Ln_Transportation_Score) - (L1_Accomodation_Score + L1_Transportation_Score) |

Again, if two of the recreational activities are selected: L1_Score_Facilities = | ((UD_L1_Accomodation_Score + UD_L1_Transportation_Score) + (2*10)) - (L1_Accomodation_Score + L1_Transportation_Score + two recreational activities score for that location) |

.

.

Ln_Score_Facilities = | ((UD_Ln_Accomodation_Score + UD Ln_Transportation_Score) + (2*10)) - (Ln_Accomodation_Score + Ln_Transportation_Score + sum of two recreational activities score for that location) |

If all five recreational activities are checked: L1_Score_Facilities = | ((UD_L1_Accomodation_Score + UD L1_Transportation_Score) + (5*10)) - (L1_Accomodation_Score + L1_Transportation_Score + sum of five recreational activities score for that location) |

.

.

Ln_Score_Facilities = | ((UD_Ln_Accomodation_Score + UD_Ln_Transportation_Score) + (2*10)) - (Ln_Accomodation_Score + Ln_Transportation_Score + sum of five recreational activities score for that location)|

Let user inputs 7 and 8 for Transportation and Accommodation facilities respectively and from the Recreational activities checkbox, he/she put 2 tick marks on Trekking and Paragliding. We are calculating score for Cox's Bazar. So, the total score becomes 7+8+10+10 = 35 from the user's side. But from TABLE IV, it is visible that Cox's Bazar, Location Id C1 has no trekking facilities and the score for Accommodation, Transportation, Paragliding contains 9, 7, 8 respectively. So, the score becomes 24. By using the above formula, we get the absolute difference between the user's choice and the existing value for that type of choices of that particular location. Here the less the difference is, the more the preference of that place will be to the user.

Budget: Budget is simply the total budget of the user for a trip. It is a crisp value. The user will also enter how luxuriously they are willing to stay and spend while they are on the trip. The system also prompts the user to specify how many days the trip will take, how many people are going on the trip and the month they are willing to visit. The month is needed to determine the ON/OFF (Peak/Off-peak) season for a particular location and the hotel accommodation cost could vary due to this.

TABLE V. BUDGET

Budget			
LID	**Spending Style**	**Season**	**Amount**
C1	Cheap	On	$70
C1	Luxurious	On	$200
C1	Cheap	Off	$50
C1	Luxurious	Off	$130

TABLE VI. SEASON

Season		
LID	**SID**	**Months**
C1	On	Oct-Feb
C1	Off	Mar-Sep

This time we are not considering the absolute difference rather considering the actual difference where negative value is considered. Let a user entered $2000 as budget preference and as additional details the user preferred to be visiting the spot for 2 days with 1 person on the month of January and also specified of luxurious staying during that visit. Each related location from the database is fetched. Then the budget score for each location is calculated by:

L1_Score_Budget = User_Budget - L1_Calculated_Cost

L2_Score_Budget = User_Budget - L2_Calculated_Cost

.

.

Ln_Score_Budget = User_Budget - Ln_Calculated_Cost

Where, L1_Calculated_Cost = Cost fetched from the database * Number of days * Number of People.

Suppose we are calculating it for Cox's Bazar that has location id C1. For C1, according to the user-defined month, it is ON season and for luxurious staying, the user needs $200 per day. So, as per the requirements of the user, the calculated cost for C1 = ($200 * 2 * 2) = $800 for 2 days along with 1 person. But, the user's preferred budget was $2000. According to the above formula, Cox's Bazar_Budget_Score = ($2000 - $800) = $1200 which is positive. In accordance to our membership function, it is Highly Affordable for that user. Another case is that, if the user had entered the budget $400, then Cox's Bazar_Budget_Score = ($400 - $800) = -$400 which is negative, that means the user is in short of budget. So, it will be less preferred for visiting that location.

Tour Type: Tour type is a variable that enables the user to specify which type of trip they wish to plan. There are 4 options to choose from: Single, Couple, Family, Friends. Suppose a location may be feasible to visit only with friends or family and not with couple or alone. That is why we have predefined scores in the database according to the feasibility of a place to visit. A place may have different scores for four types such as Friends, Family, Couple, and Alone. The value is then fuzzified and a fuzzy value is generated for this variable.

TABLE VII. TOUR TYPE

Tour Type		
LID	**TType**	**TRank**
C1	Alone	6
C1	Couple	9
C1	Friends	10
C1	Family	8

Suppose for location id C1, someone selected Tour Type as Alone. From the database, the crisp value for the Location Id C1 for type Alone is fetched. Unlike the other variable calculation this value is directly used for fuzzification. According to our membership function, it is considered as Medium. Medium indicates that it is not the best place for visiting alone.

B. Input Membership Functions:

We have used 6 fuzzy variables in the fuzzy inference system. We have used trapezoidal, triangular, and Gaussian membership functions. The different membership functions and shapes are given below:

1. Location Type: We have used trapezoidal membership function for Location Type. From 0 to 4 is High, 3 to 7 is Moderate, and from 6 to 10 is Low Similarity is used.

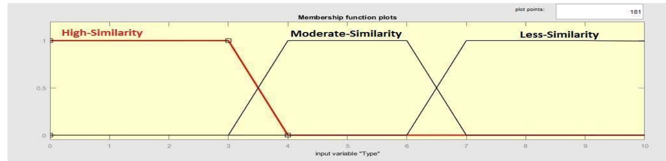

Fig. 1: Membership Function - Location Type

The generalized form of Trapezoidal Membership Function is given below

$$Trap(x;a,b,c,d)=\begin{cases} 0,\ if\ x \le a \\ \frac{x-a}{b-a},\ if\ a \le x \le b \\ 1,\ if,\ b \le x \le c \\ \frac{d-x}{d-c},\ if\ c \le x \le d \end{cases}$$

2. Beauty: We have used trapezoidal membership function for Beauty. The range is from 0 to 10. From 0 to 2 is High Similarity, 1 to 3 is Moderate Similarity, and 2 to 10 is Low Similarity. Figure 2 depicts this membership function.

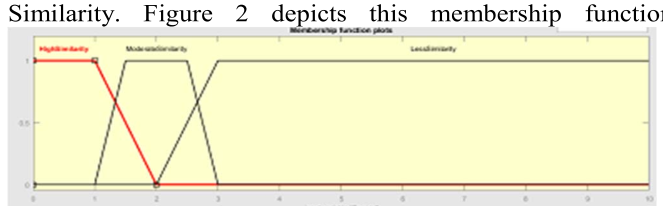

Fig. 2: Membership Function - Beauty

3. Safety: We have used Gaussian membership function for Safety. As this variable is critical, we needed the most accurate value possible. As a result, we used the Gaussian curve. The range is from 0 to 10. From 0 to 2 is Extremely Secure, 0.5 to 1.5 is Highly Secure, 2 to 5.5 is Moderately Secure, 4 to 7.5 is Highly Insecure and 6 and above is Extremely Insecure.

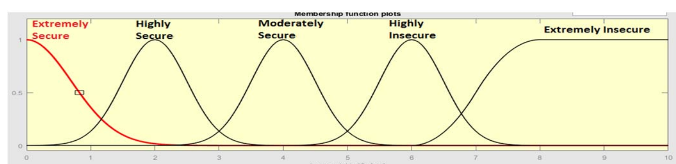

Fig. 3: Membership Function - Safety

The generalized form of Gaussian Membership Function is given (Fig. 3) below:

$$Gaussian(x;c,\sigma)= e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2}$$

4. Budget: We use triangular membership function for budget. The range is from -500 to 500. We use negative range here because while using our equation for deriving the score for user's budget, if the user's budget is lower than the calculated total cost for a location, the score becomes negative which indicates that the budget is lower. From -500 to -100 is UnAffordable, -200 to 200 is ModeratelyAffordable, and from 100 to 500 or higher is HighlyAffordable. Fig. 4 depicts this.

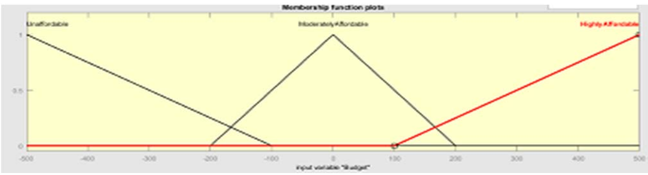

Fig. 4: Membership Function - Budget

The generalized form of Triangular Membership Function is given below:

$$Tri(x;a,b,c)=\begin{cases} 0, \text{ if } x\leq a \\ \frac{x-a}{b-a}, \text{ if } a\leq x\leq b \\ \frac{c-x}{c-b}, \text{ if } b\leq x\leq c \\ 0, \text{ if } c\leq x \end{cases}$$

5. Facilities: We have used trapezoidal membership function for Facilities. The range is from 0 to 10. From 0 to 4 is High, 3 to 7 is Medium, and 6 to 10 is Low. Fig5, shows this.

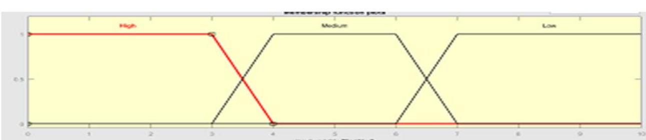

Fig. 5: Membership Function - Facilities

6. Tour Type: We have used trapezoidal membership function for Tour Type. The range is from 0 to 10. From 0 to 4 is LessMatched, 3 to 7 is Matched, and from 6 to 10 is PerfectlyMatched. Figure 6 depicts this.

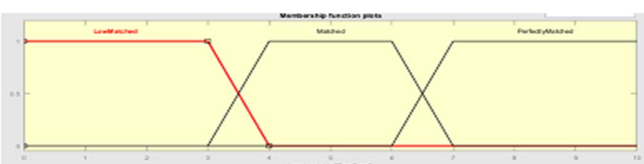

Fig. 6: Membership Function - Tour Type

C. Output Variable:

We have used six fuzzy variables in the fuzzy inference system and we have defined one fuzzy output "Score" which is calculated for each related location passed to the fuzzy controller. We infer this variable from the various IF-THEN rules in the fuzzy inference system. Five of our fuzzy variables have three levels of membership function like High, Moderate, and Low. Therefore, we have decided to assign these variables to hold a weight from 1 to 3. The other input variable, Safety, has five degrees of membership (i.e., Extremely Secure, Highly Secure, Moderately Secure, Highly Insecure, and Extremely Insecure). Therefore, this variable holds a weight from 1 to 5. Then, these six variables can hold a total weight of 20 at maximum.

The score has five levels, i.e., Very Low, Low, Medium, High, and Very High. The higher the score, the more likely it is to be suggested to the user. The range of the membership function is from 0 to 20. We have divided the levels in the following intervals: 0 to 8 is Very Low, 4 to 12 is Low, 8 to 16 is Medium, 12 to 20 is High, and 16 to 20 is Very High. Triangular membership function is used for simple calculation. The functions overlap at the midpoint of each interval.

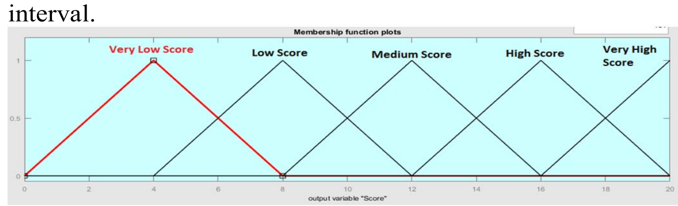

Fig. 7: Membership Function - Score

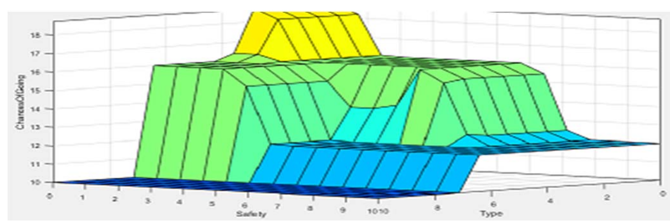

Fig. 8: Surface View of Type and Safety

When all the locations data with the respective parameters are passed on to the fuzzy controller, it uses fuzzy inference rules to calculate scores for each location. Then the system stores the locations in an ordered list grouped by their countries and the top locations are displayed to the user as output based on the score.

IV. IMPLEMENTATAION

The main purpose of the work is to develop a fuzzy inference system based recommender system that will take various inputs from user as parameters and will calculate and produce a value for each location as output. The greater the value is, the higher the preference for that location is given. The details of the system components i.e. database, fuzzy variables, inference rules etc. are given below:

A. Database:

We have used standard MySQL based database for our system. The database contains seven tables for storing necessary data for the recommendation system. These tables are- Location, Location Type, Budget, Season, Tour Type, Facilities1, and Facilities2. Fig. 9 shows the design of the complete Database Table Schema. TABLE I. stores all the tourist spots within South Asia in the database, along with their parameters, such as Beauty, Safety, etc. The ratings in this table for all the locations are predefined. We defined these values by searching various tourist spot reviews, tourist forums and by performing online surveys. TABLE II contains each location's score for different location types. Each location scores high for its own location type and low for other types. For instance, Cox's Bazar, being primarily a beach, scores higher for 'beach' type, while scoring a lower value for 'Hills'. On the other hand, since Bandarban is of type 'hills', it scores higher in that category and zero for 'beach' type, since there are no beaches in that region. This ranking system is used so that the user can perform queries with compound location types, e.g., Hills and Beaches. TABLE V contains the cost for visiting a location with different expenditure choices during different seasons. The values in this table are also predefined. We noted the total expenditure of travelers who visited these locations previously and obtained an average total cost for different spending styles. TABLE VI stores the time during which a certain tourist spot is considered as On season or Off season. This data has been gathered by contacting various travel agencies and local hotel authorities to determine which time of the year is considered On or Off season for each location in the database. TABLE VII contains the score that determines the suitability of each location for different tour types. The values are obtained in the same way as the previous one. TABLE III and IV contain facilities of accommodation and transportation and other contains recreations activities respectively.

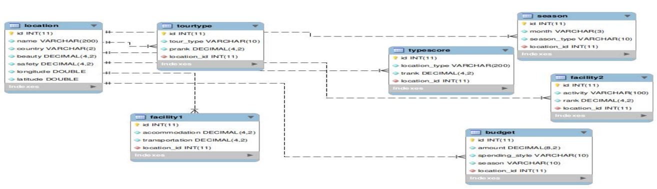

Fig. 9: Database Schema Diagram

B. The Fuzzy System:

This is the main component of the system. The fuzzy controller calculates the score of the locations. It acts as the fuzzy inference system. The values of input variables are taken from the user and all the related data are retrieved from the database. Then the system uses different input parameters for the fuzzy system. Next, the parameters and location data are passed to the fuzzy system. The fuzzy system then calculates the scores for the locations. We have used MATLAB fuzzy logic toolbox to build our fuzzy controller and Fig. 10 shows the complete diagram containing input section and output section of the system.

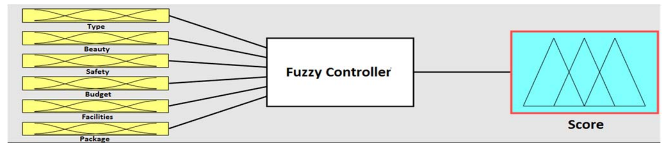

Fig. 10: The Fuzzy System

C. Fuzzy Inference Rules:

The fuzzy inference system uses six fuzzy variables to infer and calculate the score for each location. Example of a rule is "IF Location Type is HighSimilarity AND Beauty is HighSimilarity AND Safety is ExtremelySecured AND Budget is HighlyAffordable AND Facilities is High AND Tour Type is PerfectlyMatched THEN Score is Very High." In the similar way, we have generated rules for all of the possible cases.

V. Results and Discussions

We have used a sample dataset of 20 locations. We have run online surveys as well analyzed different tourist spot reviewing websites, forums, and travelling agencies to get the ratings or scores for each attribute of the locations and all other necessary data. Our system was tested on simulated data to verify the obtained results.

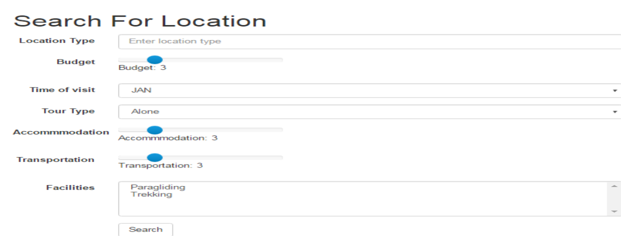

Fig.11. Sample Input Form

TABLE VIII. SAMPLE INPUT SCORE (Compound Type)

Location	Type	Bea-uty	Safe-ty	Budget	Facilities	Tour Type
Cox's Bazar	5	2	0	-300	4	9
Kuakata	10	2	3	400	7	8
Bandarban	10	1	3	300	9	7
Goa	10	2	1	-400	2	10

TABLE IX. SAMPLE OUTPUT (Compound Type)

LID	SCORE
C1	15.06
K1	11.10
B1	11.54
G1	14.38

For example, the user enters Location Type- Beach and Hills, Beauty rating 9, Safety rating 8, facilities-1 rating for accommodation and transportation 15, for facilities-2 rating, the user selected two other activities which the user wanted as recreation, budget $1780, 4 days, 4 persons, tour type Friends and preferred Luxurious staying during their travel on January. Now, the system will fetch all the necessary data that has type ‘Hills’, and ‘Beach’ in their respective location types. After that, the system will use necessary equations described in Theory Section (III) to calculate the values that will be sent to the fuzzy controller in order to fuzzify the data. Suppose, for the above set of inputs, the following values are calculated by using the formula described in Theory Section (III). Fig.11 shows a sample search query by a user.

Let the user enters Location Type- Beach this time and keeps all other entries as same as the previous one. Then the calculations will be different as well as the output score. The sample is shown below:

TABLE X.. SAMPLE INPUT SCORE (Single Type)

Locati-on	**Type**	**Bea-uty**	**Safe-ty**	**Budg-et**	**Facili-ties**	**Tour Type**
Cox’s Bazar	1	2	0	-300	4	9
Kuakat-a	1	2	3	400	7	8
Bandar-ban	10	1	3	300	9	7
Goa	1	2	1	-400	2	10

TABLE XI. SAMPLE OUTPUT (Single Type)

LID	SCORE
C1	17.60
K1	13.53
B1	9.47
G1	16.08

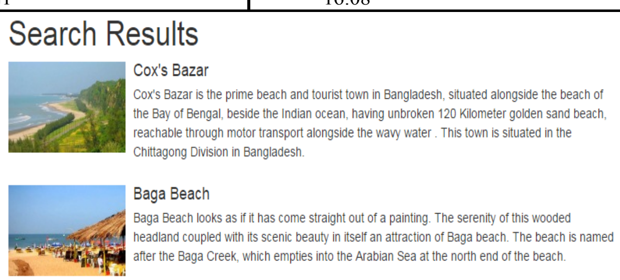

Fig.12. Sample Output Result Based on User Query

Lastly, we will provide the user the list of locations according to the score in descending order and in addition to that, we will suggest the user some other important places around that locations using the latitude and longitude value of that location. Fig. 12 gives a sample query result given by the user.

VI. CONCLUSION AND FUTURE WORK

The tourism sector is currently flourishing at a rapid rate due to people’s increasing interest in taking breaks from their regular hectic lives and embarking on a journey. Since people are inclined towards planning their trips online, the demands for tourist place recommendation systems are on the rise. However, most available recommendation systems provide outputs that are limited to user preferences and user demographic profile, or in some other cases, by using a priority system, where a priority value is assigned to each location, based on how frequently those places have been visited by others. On the other hand, our proposed recommendation system will take into account a lot of other factors such as beauty, safety, facilities available to the tourists, and type of location the tourists want to visit, and cross reference them against the user’s preference and previous visitors’ reviews and ratings and produce the optimal output for the user. As a future work we hope to improve our research by firstly introducing tables to store user preferences and demographic information; secondly enriching the database so that it is no longer limited to South Asia. We also plan to further extend the research by implementing Adaptive Neuro-Fuzzy Inference System (ANFIS) that will provide an even further personalized output for each user. Additionally, we hope to introduce a Natural Language Processing (NLP) component that should enable the system to read user reviews from various sources and perform sentiment analysis to refine user preferences.

REFERENCES

[1] Kaushik, S., Tiwari, S., Agarwal, C., & Goel, A. (2016). Ubiquitous Crowdsourcing Model for Location Recommender System. Journal of Computers, 11, 463-471. doi: 10.17706/jcp.11.6.463-471.re

[2] Makui, A., & Nikkhah, Z. (2011). Designing fuzzy expert system for creating and ranking of tourism scenarios using fuzzy AHP method. Management Science Letters, 1, 29-40. doi: 10.5267/j.msl.2010.01.007

[3] Sebastia, L., García, I., Onaindia, E., & Alvarez, C. G. (2008). 20th IEEE International Conference on Tools with Artificial Intelligence. doi: 10.1109/ictai.2008.18

[4] TripAdvsior Website: https://www.tripadvisor.com, last accessed 3rd January, 2017.

[5] Expedia Website: https://www.expedia.com, , last accessed 3rd January, 2017.

[6] Awal, M. A., Rabbi, J., Hossain, S. I., & Hashem, M. M. A. (2016, May). A hybrid approach to plan itinerary for tourists. 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), 219-223. doi: 10.1109/ICIEV.2016.7759999

[7] Hasija, H., & Chaurasia, D. (2015, September). Recommender system with web usage mining based on fuzzy c means and neural networks. 1st International Conference on Next Generation Computing Technologies (NGCT), 2015. 768-772. doi: 10.1109/NGCT.2015.7375224

[8] Wu, Q., Zhang, C. J., Gao, L., & Li, X. (2010, September). Training neural networks by electromagnetism-like mechanism algorithm for tourism arrivals forecasting. Fifth IEEE International Conference. In Bio-Inspired Computing: Theories and Applications (BIC-TA), 679-688. doi: 10.1109/BICTA.2010.5645207

[9] Indriana, M., & Hwang, C. S. (2014). Applying Neural Network Model to Hybrid Tourist Attraction Recommendations. ULTIMATICS, 6(2).

[10] Kuo, R. J., Chen, C. H., & Hwang, Y. C. (2001). An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network. Fuzzy Sets and Systems, 118(1), 21-45. doi: 10.1016/S0165-0114(98)00399-6.

[11] Yager, R. R. (2003). Fuzzy logic methods in recommender systems, Fuzzy Sets and Systems, 136(2), 133-149. doi: 10.1016/S0165-0114(02)00223-3.

[12] Zhang, Z., Lin, H., Liu, K., Wu, D., Zhang, G., & Lu, J. (2013). A hybrid fuzzy-based personalized recommender system for telecom products/services, Information Sciences, 235, 117-129. doi: 10.1016/j.ins.2013.01.025..