VISVESVARAYA TECHNOLOGICAL UNIVERSITY,

JNANA SANGAMA, BELGAUM - 590014

A SRS project report on

Personalized Tourism Itinerary Recommender System

Submitted in partial evaluation of the 8th semester project progress review-1

Submitted by:

1PI13CS047	Chiranth Gopal K
1PI13CS062	Gourish Hebbar

Under the guidance of

Dr. B Narsing Rao
Professor, PESIT

Jan – May 2016

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

PES INSTITUTE OF TECHNOLOGY,

(AN AUTONOMOUS INSTITUTE UNDER VTU, BELGAUM AND UGC, NEW DELHI)

100FT RING ROAD, BSK 3RD STAGE, BENGALURU - 560085

Abstract:

When we want to go visiting places we are new to, we would not know what places to go and which are good ones that we would like. So, we are building a web app, in which the user will first login and enter his preferences. Each user, once he logs in, will have option to tell what his duration of trip will be, i.e, the dates, and for each day, the time he wishes to go visiting places, the start/end point of the trip and also the type of places he wishes to see that day. Then using this information, if the user has already visited certain places and has given ratings to that place, then using that information, using the rating information for places, average time spent at place, location information stored in the database, we will generate an itinerary for the user. It will contain a map, with the recommended places to visit, in the shortest time possible between places, with the start point the user selects as the destination point as well. So, with this, the user can get a good idea as to how he could effectively use his time to visit the places he likes without having to scout through a bunch of different websites and looking up their locations.

TABLE OF CONTENTS

1. Introduction:

Current landscape of tourism travel in India is that if a person wants to visit a place, he can find a few websites which will allow him to search for places around or give recommendations on the places based on his preferences. But say if we have some time constraints, we do not have something that would give us an entire itinerary which would consider our preferences, time constraints, the time to spend at the place etc. and give a map with our itinerary. The user may not know much about the locations of the places, so he can be rest assured that the path we suggest is the most optimal one, thus saving him precious time as well as money. Since the user may also not know the time slot that the place is open to visiting, it is difficult to manually search details about each place and then plan accordingly, or else he may go and visit during wrong hours. So we see that these problems are actually pretty common, but they have not yet been properly addressed. So we set out to solve it, by building a web app.

1.1 Problem definition:

To provide personalized itinerary for attractions based on user location and preferences.

1.2 Generic proposed solution:

These are the detailed steps of our solution:

1. Database: We are building a database, which is populated by scraping the data from multiple sites like TripAdvisor, Google, which we collect and create a database using MySQL. Since each site will have it's own format of storing data, we are trying to create modules, which are designed specifically for the particular site, but all of the modules give out standard format of data to the database. We are also increasing the level of abstraction, by scraping using code, and making a generic data parser, rather than using tools and manually scrape the data.

2. Recommender engine: We are using the rating data of the place from different sites, number of reviews, data from the user's previous places and his ratings, his current preference, his location and all of this will be assigned weights. Then, we take the weighted average of these scores, and calculate the scores for each place. Then based on the proximity of places, the assigned scores and the time slot that the user is free, we will use a variant of Travelling Salesman algorithm where we try to maximize the weights and create an itinerary, which can be edited by the user.

3. Front end: The UI will consist of a web app, with login to authenticate the user. Then, we take the following inputs from the user:
   - Dates of travel
   - Time of availability for the day
   - Start/end location of travel
   - Category of places

Then, we will take inputs and generate a few recommendations, and using those we generate an itinerary, which will be shown on the map, with the list of places on the side. The places not included in the itinerary are also shown, so the user can add/remove places from the itinerary and new itineraries will be generated on the fly.

We can say the project would be complete, when our web app is ready, which is by April 14 with above functionalities and features included.

Signature of the guide

2. Literature survey:

Most mainstream tourism web sites that exist today, have certain main functionalities. Below, we mention those main things that they are capable of doing:

- MakeMyTrip, TripAdvisor:

  These sites offer a lot of functionalities, like being able to book hotels, air tickets and restaurants. The differences are that, MakeMyTrip is more of travel oriented site, with more options to book travel options like bus and trains, whereas TripAdvisor contains exhaustive knowledge related to places. TripAdvisor can be used to search for places nearby, based on preferences of user, but it is not possible to create an itinerary of places to visit.

- Google Now :

  Google generates recommendations automatically on places nearby once we are in a given location. We can also search based on the type of places we want to see, and it will show places sorted in increasing order of distances from the our current location. There is no option to create an itinerary of places that we would want to see

- Inspirock:

  Inspirock takes a place as input from the user, and generates an itinerary of places, with option to add or remove places. This is the most similar existing solution, but there are however certain differences, them being:

  There is no option to set the start and end time of the trip for the day

  User cannot extend the time which is fixed at 6:00 PM in the evening, so rules out places to visit after sunset

  The start and end points are different, and cannot be changed by the user

Once an itinerary is created, there is no option to get recommendations to restaurants closer to the place which the user is currently in if the users want to have a break for eating

2.1 References:

1. Location-Aware Recommendation Systems by María del Carmen Rodríguez-Hernández, Sergio Ilarri, Raquel Trillo-Lado, Ramón Hermoso - Article 2015

This paper provides a survey on location-aware recommender systems in scenarios involving mobile computing. It describes the fundamentals of recommendation systems, shows the most relevant existing approaches for location aware systems. Then it talks about the current applications of location aware recommender systems in different domains. From this paper, we got an idea of how recommender systems are used in various cases, and how it applies in our case of tourism. This prompted us to think in terms of what we could do and how we could leverage the abundance of location data to build a good recommender system.

2. The Use of Machine Learning Algorithms in Recommender Systems: A Systematic Review by Ivens Portugal, Paulo Alencar, Donald Cowan - arXiv 2015

This paper presents a review of the literature that analyzes the use of machine learning algorithms in recommender systems and identifies the research opportunities in the field of software engineering research. Along with comparing the use of algorithms, it also compares the usage of recommender systems in different domains. From this, we learnt that among the papers that the authors have reviewed on recommender systems, the two most used algorithms were Bayesian and Decision trees. The authors think that this is because of the simplicity of the

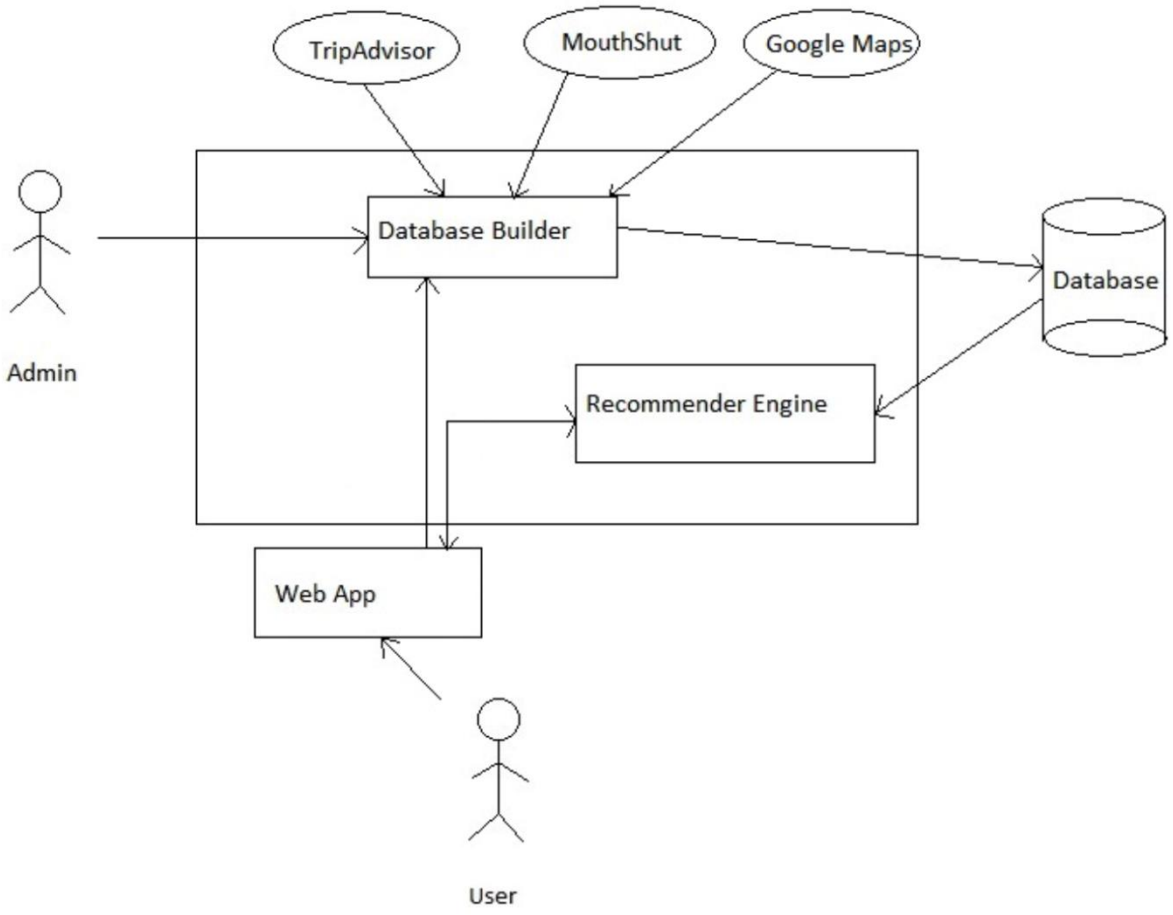
algorithms, i.e, since they are less computationally expensive. Another important thing that we found was that among the papers, tourism was among the last in the list of domains ordered by the number of papers on them. So it was an added incentive for us that we were choosing a field where there is not much research done before.

3. Web application for recommending personalised mobile tourist routes by D. Gavalas, M. Kenteris, C. Konstantopoulos, G. Pantizou - IET Softw., 2012

This paper deals with the problem of creating personalized recommendations for daily sightseeing itineraries for tourists visiting places. Their approach considers a few selected places of interest that a traveler would potentially want to visit and derives a near optimal itinerary solution for each day of visit and the places of potential interest are selected based on stated or implied preferences. Their method enables planning of customized, each day based, personalized tourist itineraries considering user preferences, time available for visiting sights on a daily basis, opening days of sights and average visiting times, using heuristics. This was the most relevant paper for our reference, as the concept of creating itinerary is same, and also many of the features. But they have not given their algorithm, so we will have to develop our own.

3. System Requirement Specification:

3.1 High level architecture diagram



3.2 Hardware Requirements

- Server - laptop
- Client device - laptop, phone

Since we are not yet planning to host our web app online after completion, our hardware requirements, i.e., server and client, are both our own personal devices.

3.3 Software Requirements

Programming Languages :

- Front end:
  - Javascript, HTML, CSS, Bootstrap

    Our familiarity with Javascript and HTML, and their robust usage around internet, led to us deciding on them. Bootstrap is chosen for a more responsive experience
- Back end:
  - Server side: PHP

    PHP was chosen as the backend language of choice
- Data extraction:
  - Python

    Python has vast amounts of support and libraries. Some of the libraries that we are using are mentioned below:
    - beautifulsoup: Data Extraction
    - selenium: Parsing web sites
    - googleplaces: Getting latitude longitude data from place name
    - pickle: Storing and retrieving intermediate data
    - MySQLdb: To connect with MySQL database

- Database:
  - MySQL

3.4 Functional Requirements and Test Cases

- FR1 Login page
  - Description: Login page for user account access.
  - Evaluation-T1: Manual. We have to create an account and try to login using that account.
- FR2 Set trip dates
  - Description: The dates on which the user intends to go on the trip.
  - Evaluation-T2: Manual. We have to enter dates of the trip.
- FR3 Set location
  - Description: The place in which user chooses to go around sightseeing other places.
  - Evaluation-T3: Manual. Location is choses by us, so have to manually choose in the map.
- FR4 Set category preferences
  - Description: Set the type of places the user wants to see.
  - Evaluation-T4: Manual. Category should be chosen manually.
- FR5 A
  - Description: Set the start and end time for the day.
  - Evaluation-T5: Manual. Time slots need to be chosen manually.
- FR6 Generate itinerary
  - Description: Create the itinerary on google maps and show user a list of the recommended places.

- Evaluation-T6: Manual. See if the generated itinerary is properly displayed on the map

- FR7 Add or remove places from recommended list
  - Description: Give user option to manually add or remove places from the itinerary generated.
  - Evaluation-T7: Manual. We have to test by manually adding and removing places.

- UI1 Provide login page
- UI2 Data input page
- UI3 Map generation/places recommender page

3.5 Non Functional Requirements

- NF1 Quick data access
  - Description: When retrieving the data from the database, fast retrieval helps in generating recommendations quicker
- NF2 Reliability - Database has ACID properties (MySQL)
  - Description: MySQL comes with ACID properties of Atomicity, Concurrency, Isolation and Durability
- NF3 Optimal path generated
  - Description: Since the recommendations generated vary from user to user, the path in the itinerary generated will also vary. But we will be generating the optimal path among those recommendations

3.6 Constraints

- Data is not available. So we have to scrape the websites and extract data from them
- Some sites have dynamic content generation, i.e., the elements on the site are created by javascript functions on some action by the user. So we need to use a web automation tool to emulate the actions as it would be made by the user, and then grab the data
- As the recommendations that are generated will be different for different users based on their preferences and location, the generated itinerary will be different, thus making it harder for evaluation
- Generating shortest path is a NP-hard problem and hence it would take time to generate itineraries on the fly based on user input

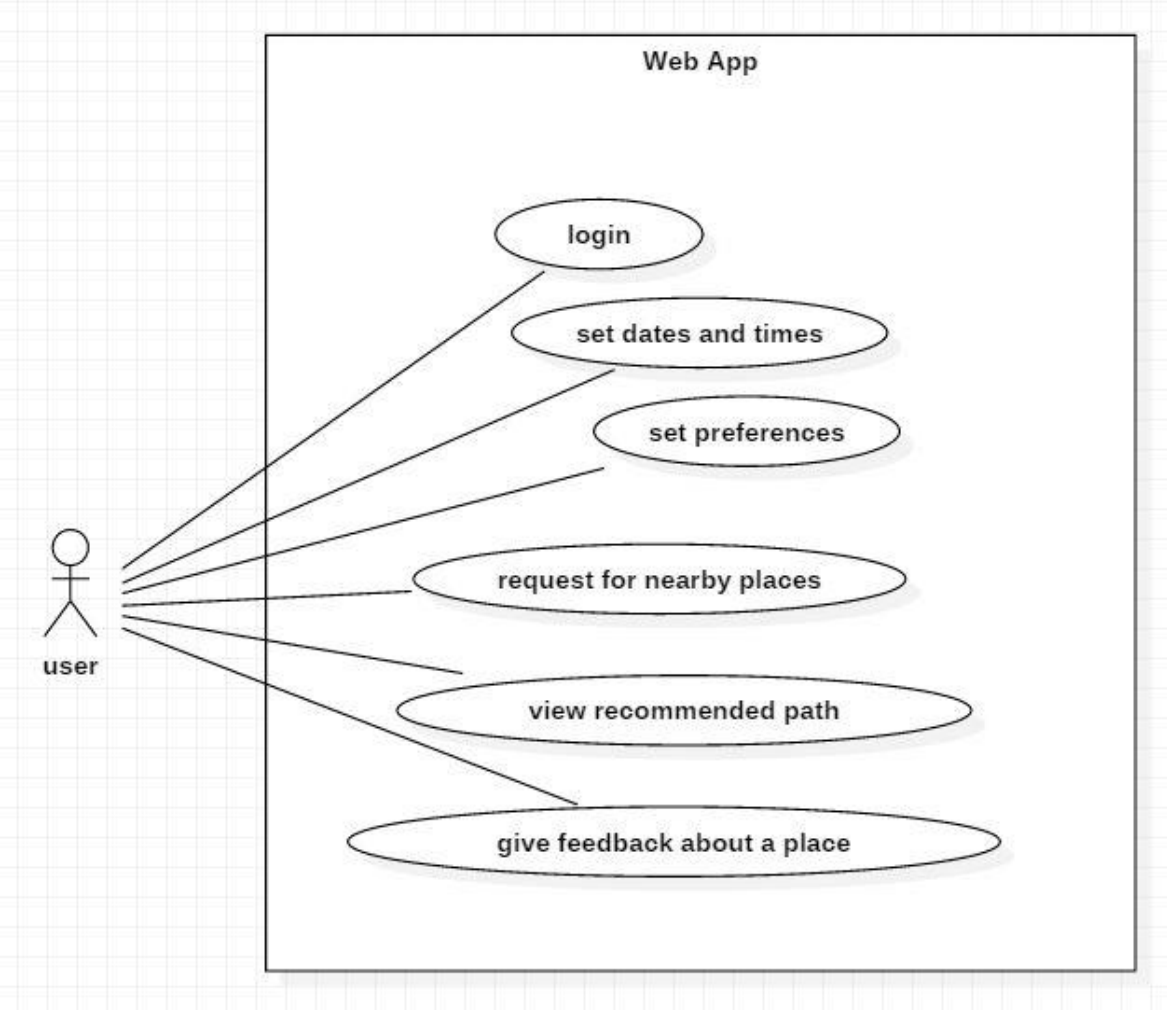
3.7 Requirement Traceability Matrix (RTM)

Requirement ID	Requirement Type	Requirement Description	Test scenario	Results	Comments	
FR1	Functional	Login page for user account				
FR2	Functional	Setting dates for the trip				
FR3	Functional	Setting location of the trip				
FR4	Functional	Setting category preferences				
FR5	Functional	Setting time slots for the day				
FR6	Functional	Generating itinerary				
FR7	Functional	Adding/removing places from list				
UI1	Functional	Providing login page				
UI2	Functional	Data input page				
UI3	Functional	Map generation page				
NF1	Non Functional	Quick data access				
NF2	Non Functional	Reliability				
NF3	Non Functional	Optimal path generation				

3.8 Use case scenarios

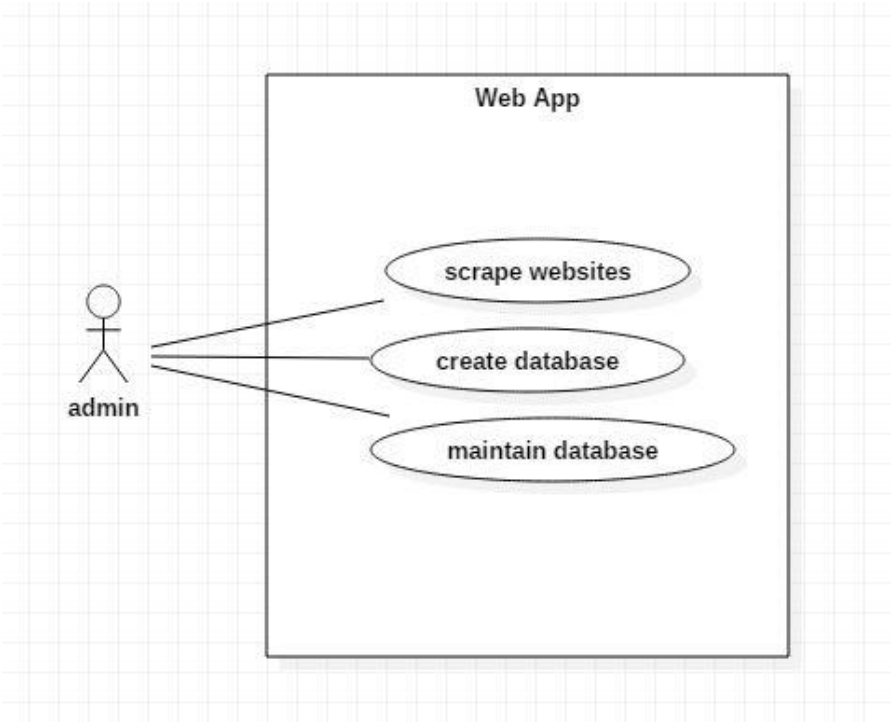
- User

  The following diagram denotes the high level use cases of the user's interaction with the web app.



- Admin

  The following diagram denotes the high level use cases of the admin's interaction with the web app's database.



3.9 Gantt Chart