BACKPACKER'S INDIA

TEAM 6
AVINASH SANKAR GANDHI RAJAN
KARTHIKEYAN VIJAYAKUMAR
PRAKASH PONNUSWAMY
RUFUS DEEPAN RAVICHANDRAN
CONTENTS

INTRODUCTION ......................................................................................... 2
BUSINESS MODEL .................................................................................. 2
COMPETITOR ANALYSIS ............................................................... 3
NOVELTY .......................................................................................... 3
USER SCENARIOS ............................................................................. 5
  Scenario #1 .................................................................................... 5
  Scenario #2 .................................................................................... 6
  Scenario #3 .................................................................................... 7
SYSTEM ARCHITECTURE .............................................................. 8
API FUNCTIONALITIES ...................................................................... 8
ANALYTICS ...................................................................................... 9
  Twitter Analysis ......................................................................... 9
  Analysis of review comments: ...................................................... 10
  Sentiment classification phases ............................................... 11
FUTURE EXTENSIONS ........................................................................ 11
PROJECT TIMELINE .......................................................................... 12
TEAM MEMBER CONTRIBUTIONS .................................................. 13
REFERENCES ..................................................................................... 14
INTRODUCTION

“Backpacker’s India” is a website that aims at addressing the necessities of backpackers who travel to India. As per Wikipedia, backpacking is “a form of low cost, independent international travel”. Any backpacker travelling to another country is looking at soaking in the culture of that country unlike casual travelers. Our website provides backpackers this experience during their trip to India through recommendations on what meditation/yoga/Ayurveda centers and temples are a must visit, where to stay such as local hostels, how to get around using local transport, what local people suggest as a must see in that city. The other informative feature of the website is its availability of the wiki search field with which the user will be able to find the links related to the search term. Some of the other important features of Backpacker’s India include current weather display of a city, live currency converter, dynamic display of twitter feeds related to a city, wiki search and the option to view the website in any of the google translate languages.

BUSINESS MODEL

The following stakeholders will contribute towards the revenue of “Backpacker’s India”:

- **Travel Agencies** – All the travel agencies including the local private bus agencies, cab agencies like Redbus.com, fasttrackcalltaxi.in that are being listed in the website when people search for places in the “How to move around” page, can contribute to the revenue for the website. We will collect a fee from all those agencies who wish to get them listed in the results.

- **Hostels** – Since most of the backpackers would love to stay in an inexpensive place when they visit different cities of India, those budget hostels and hotels become good avenues for revenue generation. We can collect a nominal fee from those hostels for them to get listed in the “where to stay” page. There is also a possibility of collecting commission fee from hostels for every referral that gets directed from our website.

- **Advertisements** - Online advertising is another way to generate revenue. Since a lot of users come searching for information, they tend to click on the advertisements that we put up.
  
  - **Google Ads** - We can use Google AdSense to have user specific ads shown on our website, and generate revenue if someone clicks on it.
  
  - **Banner Ads** - We can also use banner ads to earn additional revenue from our website. We would like to make it a free to use website as people would not want to pay to visit an informational website.
COMPETITOR ANALYSIS

The unique features that Backpacker’s India provides over the other websites are the options for local transport including cost, timing and fare, local user recommended attractions, hostel recommendations based on user reviews and comments and twitter sentiment analysis to determine the overall sentiment towards cities in India.

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>BACKPACKER’S INDIA</th>
<th>INDIANBACKPACKERS</th>
<th>THEBACKPACKERTOURIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>PUBLIC/LOCAL TRANSPORT</td>
<td>✔️</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>LOCAL ATTRACTION</td>
<td>✔️</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>MEDITATION CENTRES, AYURVEDIC CENTRES, TEMPLES</td>
<td>✔️</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>TRAVEL BLOGS/FORUMS</td>
<td>✗</td>
<td>✔️</td>
<td>✗</td>
</tr>
<tr>
<td>TRAVEL COMPANION</td>
<td>✗</td>
<td>✔️</td>
<td>✗</td>
</tr>
<tr>
<td>HOSTEL RECOMMENDATIONS</td>
<td>✔️</td>
<td>✔️</td>
<td>✗</td>
</tr>
<tr>
<td>TWITTER SENTIMENT ANALYSIS</td>
<td>✔️</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

NOVELTY

“Backpacker’s India” provides 3 novelties which other backpacker websites do not provide

1. **User recommended suggestions and tips about places to visit in the area of interest**
   Many travel websites mention all the major tourist places in a location but fail to mention the local delicacies or hidden gems of the location. Our website provides information on the local delicacies through other users’ comments and tips on what to try out in the area of interest. Moreover, a backpacker is interested in the untrodden path. We pacify this interest by mentioning
the local Ayurveda centers, meditation/yoga centers, famous temples and local cuisines that enable them to soak in the culture of India.

2. Local transport options along with their time and fare
   A major concern for backpackers when they travel to India is how to get around. Unlike the casual traveler, backpackers look at using the local transportation available in the city that they visit. We alleviate this concern by providing them local transportation information within a city as well as across cities (getting from one city to another). All that the user has to do is enter the source/start address and destination/end address. A list of different modes of transport—buses, trains, flights are provided along with their timings, journey time and current fare in US dollars. Another unique feature is the use of maps to provide users to identify the route to be taken in case they decide to walk to the destination. US dollars has been used as the standard currency since backpackers from around the world can relate to US dollars. An option to convert US dollars to the required currency is also provided with the current international exchange rates.

3. Recommendations on local hostels based on aggregated user sentiments
   Our website only lists hostels in the local area that are cheap and those that have user comments. Cheap hostels do not always translate to being safe, well maintained and clean. We have provided a list of hostels in each city with aggregated user sentiments towards them. A backpacker can check out a particular hostel that he is interested in and see whether people who have stayed previously in that hostel have had a positive, negative or neutral attitude towards that hostel. This will greatly help a backpacker make decisions on which hostel to stay when they visit the city.
USER SCENARIOS

Scenario #1
A backpacker who has little knowledge of different cities in India visits this website. He/she can refer the analysis in the homepage to explore few interesting cities. He selects the city that he likes in the homepage. Once he selects the city, he will have information about famous places nearby, location information along with tweets related to backpackers in India. He can get more information about that city and hence, he could decide his city and few important places that make his trip valuable.

After he gets an idea of particular city, he might be interested in exploring Indian culture. He can see nearby Ayurveda center, meditation center, temples as well as local cuisines. He can also learn more about the location or any of the centers provided above in wiki using wiki search API. This helps him get the complete picture of the place that he planned to visit.
Scenario #2
This scenario considers a backpacker who decided a city and plans for a local travel. He/she can use our site and get the granular level details on how to reach his destination. We have our registered travel agents whose website links will be available for the backpacker to book his itinerary.

He can use the currency converter to calculate the accurate expense that it might take and our site will provide total expense as well as individual expenses from the start till his/her destination. He/she can join with other backpackers or discuss with others to get more information using our simple chat room.

How to get around India?
Chennai, Tamil Nadu, India → Goa, India

Train from Chennai to Madras (MAA)

<table>
<thead>
<tr>
<th>AGENCY</th>
<th>FROM</th>
<th>TO</th>
<th>FREQ</th>
<th>DURATION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chennai Fort</td>
<td>Tiruvelam</td>
<td>every 15 minutes</td>
<td>34 minutes</td>
</tr>
</tbody>
</table>

Schedules at: [SR indianrailways.gov.in](http://sr.indianrailways.gov.in)

Prefer a taxi instead? Chennai to Madras (MAA) will take 35 minutes (11.4 miles) ONE WAY FROM $6 USD
He can use the currency converter to calculate the accurate expense that it might take and our site will provide total expense as well as individual expenses from the start till his/her destination. He/she can join with other backpacker or discuss with others to get more information using our simple chat room.

**Scenario #3**
As an added value, a backpacker who decided his/her city can use our website to find best and safe hostel to stay in. We have the analysis report to choose the best hostel in that city. We have the address and overall rating of that hostel/guest house which will aid the backpacker to decide his accommodation.
**SYSTEM ARCHITECTURE**

The system architecture for *Backpacker’s India* is straightforward. We used Eleven APIs for the website and each API is called through a combination of PHP and JavaScript. We have used Apache as the application’s web server. The database that stores all information about the different places and the details of details is hosted on MySQL. The website is housed on Amazon’s EC2 cloud instance.

We have implemented a web crawler that scrapes information from three different hostel booking websites that provides reviews and safety details about budget hostels present in different cities of India.

**API FUNCTIONALITIES**

A total of Eleven APIs have been used in this web application. Each API assists in one way or the other to provide meaningful information to the user who visits the website with the aim of getting to know more about the place he wishes to travel in India.

Following table would give a better picture of the functionalities of all the Eleven APIs

<table>
<thead>
<tr>
<th>APIs</th>
<th>Functionalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Maps</td>
<td>Provides maps of temples, local cuisines and yoga/meditation centers to the backpackers</td>
</tr>
<tr>
<td>Wikipedia</td>
<td>Provides information from Wikipedia to users when searched</td>
</tr>
<tr>
<td><strong>Facebook</strong></td>
<td>To share the webpage URL on the user’s Facebook page</td>
</tr>
<tr>
<td><strong>Flickr</strong></td>
<td>Displays images of the city the backpacker views</td>
</tr>
<tr>
<td><strong>Google Places</strong></td>
<td>To provide the Google’s autocomplete feature when searching for places to navigate to</td>
</tr>
<tr>
<td><strong>Twitter</strong></td>
<td>Shows tweets of fellow backpackers around the world</td>
</tr>
<tr>
<td><strong>Rome2rio</strong></td>
<td>Helps discover how to get anywhere in India by searching plane, train, bus, car and showing routes and estimated cost</td>
</tr>
<tr>
<td><strong>Foursquare</strong></td>
<td>To display the local attractions of the city the user is interested in</td>
</tr>
<tr>
<td><strong>Chat API</strong></td>
<td>Enable chatting with fellow backpackers who are online and plan travel together</td>
</tr>
<tr>
<td><strong>Weather API</strong></td>
<td>Displays the weather in the city for the week</td>
</tr>
<tr>
<td><strong>Currency Converter API</strong></td>
<td>Assists user in converting local currency to their own currency with ease</td>
</tr>
</tbody>
</table>

**ANALYTICS**

Our analytics has 2 parts to it.

**Twitter Analysis**

Twitter feeds are collected with different combinations of hash tags such as city names, few keywords like stay, accommodation and so on. We used SentiWordNet analysis to classify all the tweets collected from these hash tags. These tweets are stored in database using MYSQL. This is provided as an input directly to SentiWordNet algorithm. It uses AFINN-111 lexical word to classify tweets. Tweets are classified as positive, neutral and negative. Based on the count of this result, we suggest users to select cities particularly when user is willing to travel to any such city in India.
Analysis of review comments:
We had more than 2 websites which carried user comments and rating of these hostels. Also, they have provided description of how their experience was after staying in those hostels. We scraped all these static user comments along with the location and their overall rating. We then used these comments together to analyze the overall performance along with the safety measure of these hostels.

AFINN – 111:
This is a lexical dictionary that has 2477 words in it and it scores in the range of -5 (negative) to +5 (positive). It is used by Sentiwordnet algorithm to give scores for each word. This dictionary has predefined scores attached to it. WordNet distinguishes between nouns, verbs, adjectives and adverbs. The assigned scores are within the range 0 and 1, 0 being the lowest and 1 the highest.

When a word is not available in the dictionary, the algorithm learns the nature of the score based on the below formula,

$$S_W = \frac{\sum_{t} S_t(w \in t)}{n}$$
Where, \( n \) is the number of tweets that contain the word, \( S_t \) is the sentiment score of the tweet that contains the word \( W \). Hence, the higher is the number of tweets as inputs, greater is the precision of the sentiment for the new score.

**Sentiment classification phases**

![Diagram of sentiment classification phases]

- **Tokenization and speech tagging:**
  This splits the text into very simple tokens such as numbers, punctuation and words of different types. Speech tagging produces a tag as an annotation based on the role of each word in the tweet.

- **Word sense disambiguation:**
  This is the method that is used to identify the best context of words being used in the tweets. We have implemented this using NLK, a python library for NLP (natural language processing).

- **Sentiment Orientation:**
  The positive and negative scores for each term found in a tweet are summed separately to get 2 scores. However, we considered average score of tweets. This is calculated by the average scores of positive and negative in that tweet. Finally, sentiment of a tweet is determined based on the higher value between \( S^+ \) and \( S^- \):

  \[
  s_t = \begin{cases} 
  \text{positive if } s^+ > s^- \\
  \text{negative if } s^+ \leq s^- 
  \end{cases}
  \]

**FUTURE EXTENSIONS**

- **Expand scalability to all cities and states in India**
  The website currently supports the top 25 cities that backpacker’s prefer to visit in India. In the future, all cities and states in India can be incorporated into the website. Other unknown local attractions can also be added to the list that the website already hosts.

- **Forums/travelogues for personal user experiences**
  Moving forward forums based on cities in India can be incorporated into the website. Backpackers can write blogs on their travel experience. Other backpackers can review these travelogues and gain insight on what a particular backpacker should do or should not do when they travel to India.
At the same time, forums will help backpackers who travel to India share their thoughts and views about a particular city.

- **Integrate Amazon to recommend products**
  Based on the city that a particular backpacker plans to visit, essential items required for travel can be recommended to the user. For example, a backpacker travelling to a hot and humid city such as Chennai can be recommended linen clothes, hats, on-the-go water filters and sunscreen lotions.

- **Location recommendations based on user activity**
  The cookies that are collected when a user visits the website can provide valuable information about that user. Based on the cities and type of local attractions that user frequently checks out on the website similar cities and places of interest can be provided to the user upfront when he visits the website again.

### PROJECT TIMELINE

<table>
<thead>
<tr>
<th>Task Name</th>
<th>Duration</th>
<th>Start</th>
<th>Finish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backpacker's India Project</td>
<td>74 days</td>
<td>Sat 2/1/14</td>
<td>Wed 5/14/14</td>
</tr>
<tr>
<td>Project Proposal</td>
<td>2 days</td>
<td>Sat 2/1/14</td>
<td>Mon 2/3/14</td>
</tr>
<tr>
<td>Mentor Meeting</td>
<td>1 day</td>
<td>Wed 2/12/14</td>
<td>Wed 2/12/14</td>
</tr>
<tr>
<td>Research - Identification of data sources/APIs</td>
<td>6 days</td>
<td>Sat 2/15/14</td>
<td>Fri 2/21/14</td>
</tr>
<tr>
<td>Website layout and database schema design</td>
<td>8 days</td>
<td>Wed 2/26/14</td>
<td>Fri 3/7/14</td>
</tr>
<tr>
<td>Collection of 5 APIs for demo 1.0</td>
<td>9 days</td>
<td>Tue 3/11/14</td>
<td>Fri 3/21/14</td>
</tr>
<tr>
<td>Implementation of website design</td>
<td>5 days</td>
<td>Mon 3/24/14</td>
<td>Fri 3/28/14</td>
</tr>
<tr>
<td>Integration of APIs into the website</td>
<td>4 days</td>
<td>Sun 3/30/14</td>
<td>Wed 4/2/14</td>
</tr>
<tr>
<td>Testing the website</td>
<td>2 days</td>
<td>Fri 4/4/14</td>
<td>Mon 4/7/14</td>
</tr>
<tr>
<td>Version 1.0 demo</td>
<td>1 day</td>
<td>Mon 4/7/14</td>
<td>Mon 4/7/14</td>
</tr>
<tr>
<td>Redesign of website after review</td>
<td>3 days</td>
<td>Tue 4/8/14</td>
<td>Thu 4/10/14</td>
</tr>
<tr>
<td>Addition of additional APIs to the website</td>
<td>5 days</td>
<td>Fri 4/11/14</td>
<td>Thu 4/17/14</td>
</tr>
<tr>
<td>Analytics - user review analysis on hostels</td>
<td>4 days</td>
<td>Fri 4/18/14</td>
<td>Wed 4/23/14</td>
</tr>
<tr>
<td>Analytics - twitter feed sentiment analysis</td>
<td>4 days</td>
<td>Thu 4/24/14</td>
<td>Tue 4/29/14</td>
</tr>
<tr>
<td>Integration of analytics components to the website</td>
<td>3 days</td>
<td>Wed 4/30/14</td>
<td>Fri 5/2/14</td>
</tr>
<tr>
<td>Testing the website</td>
<td>3 days</td>
<td>Thu 5/1/14</td>
<td>Mon 5/5/14</td>
</tr>
<tr>
<td>Version 2.0 demo</td>
<td>1 day</td>
<td>Mon 5/5/14</td>
<td>Mon 5/5/14</td>
</tr>
<tr>
<td>Project report</td>
<td>3 days</td>
<td>Sat 5/10/14</td>
<td>Tue 5/13/14</td>
</tr>
<tr>
<td>Project report submission</td>
<td>1 day</td>
<td>Wed 5/14/14</td>
<td>Wed 5/14/14</td>
</tr>
<tr>
<td>Team Member</td>
<td>Contribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Avinash Sankar Gandhi Rajan  | • Acted as a business analyst, actively developed business scenarios and identified the novelty for our website  
• Identified the right location for integrating all APIs and place holders in each page  
• Integrated Google maps API, Flickr API |
| Karthikeyan Vijayakumar     | • Played a major role in scraping static contents, sentiment analysis and database management  
• Designed the website layout that suits an ideal travel website  
• SentiWordNet analysis, Chat functionality and Google Places API |
| Prakash PonnuSwamy           | • Drove project to success from start, identified revenue models and novelty factors in website  
• Created Ec2 instance single handedly, installed essential software in Ec2 and managed database mapping with front end  
• Implemented Twitter API, Currency converter API and Weather API |
| Rufus Deepan Ravichandran   | • Web developer of the project, designed the front end, analyzed different websites to identify relevant APIs  
• Implemented the website layout and all individual pages in the website  
• Implemented Rome2Rio API, Foursquare API and Wiki API |
REFERENCES


