**Intelligent Web Assignment**

team name: unkonwn

member: Pengyuan Zhao, Likang Cao, Shangshu Lu

# 1 Introduction

introduction

# 2 For each task in section 1

## 2.1 Querying the social Web

### 2.1.1

**Issues:**

The task of this part is getting relevant messages through Twitter according to queries that users input. Based on the requirement, it can limit the query to a specific geographic region(latitude/longitude). Moreover, a demonstration of people who retweeted each message must be achieved. And the most difficult thing is to display who retweets the tweets and provide a hyperlink for the message.

**Design and its Motivations:**

This part we use Search Tweets api to get useful information according to user’s input. According to the requirement, we should use keyword and geographic region(like latitude and longitude) to get the discussions. Search Tweets api provides two useful parameters which are query and coordinates, that can specify the requirement.

**Requirements:**

We can perfectly satisfy the requirements, which are showing the relevant information and displaying the users who retweeted the messages.

**Limitations:**

The extension of this part is that the web may return a graph that shows the same venues these user both visited

### 2.1.2a

**Issues:**

The objective of this part is extracting tweets from specific users and analyze the high frequent keywords existed in the last few days. The keywords will be sorted depending on the total amount. The most frequent keywords Y used in the last days Z must be displayed. Moreover, the keywords must common to all users.

The first challenge of this part is that the tweets from users will contain a lot of punctuations and blank. This part must be considered, otherwise there would be a plenty of illegal keywords come out. Secondly, the high frequency keywords may not exist in a specific user's tweets, which should be deleted from the whole wordlist to satisfy the requirement of "common to all users".

**Design and its Motivations:**

First of all, we need extract tweets from a specific users. The days and users' screen\_name can be controlled by the search/tweets API. Following by token the tweets and eliminate the punctuations and blanks, we store them into arrays.

After we repeat this step for a few times(which are limited by the number of users entered), we can get a series of arrays which are contained all text of tweets we get from the API. We can combine them all, and divide it into individual words. Next, we could calculate the frequency of each words and sort them in the order. According from the keywords number we entered, we can put these specified high frequency keywords into an array. At last, we compare this array with the arrays of each users tweets, and get the date of each keywords for specific users.

**Requirements:**

we can return the correct information to the frontpage, including the screen\_name and the keywords we entered.

**Limitations:**

The process of getting and handling data is quite complex. All parts of implements are designed in the same function, there is lacking in several interface and encapsulation to control these functions.

### 2.1.2b

**Issues:**

The aim of this part is to find venues the user visited in last several days, the main strategy to solve this problem is shown below: Firstly, I used the search/tweets API to get the user’s tweets with Swarmapp check-in messages; then, find use split() to find the check-in id from user’s tweets; after that,use foursquare checkins resolve API(https://api.foursquare.com/v2/checkins/resolve) to get venue name through the check-in id we got.

There are so many Issues I met when I was doing this problem; firstly, I need to look for a correctly API from both twitter and foursquare API documentations to get desire date. Secondly，I had to analysis data I got to retrieve the desire information(ie. Check-in Id),,after that, the largest issue is that some users cannot find through Twitter and foursquare API, so that it makes a large challenge during testing the system.

**Design and its Motivations:**

For overcoming the problem of searching user’s tweets with check-in information, I used search/tweets API swarmapp/com/c/ as query and set parameter “from ”be a given user, and I used setDate() to get the days we need.

For finding out all the user’s tweets with check-in message, I split ‘expanded\_url’ properties of tweets to get check-in ids because it is easy to find that all the check in messages are shown by the format swarmapp/com/c/check-in ID,Afterthat,I used foursquare API 'https://api.foursquare.com/v2/checkins/resolve' to get venues name.

For storing the information, I built two tables in the database which are named “users Information”(which used to store “user name”, ”user id”, ”user location” and “user’s description”) and ”locationVisited”(which used to store user id, venue id and venue name) ,I design them with the concept of relational database, I used “user id” in the table “usersInformation” as the foreign key in the table “locationVisited”.

**Requirements:**

**Limitations:**

Firstly the system is not so robust and always suffers from all kinds of issues, such as some users cannot find through this system.Secondly, the system is not so efficiency and always cannot retrieve some user when processing.Finally, the database is not so efficiency when accessing data.

### 2.1.2c

**Issues:**

This part is to get a list of user by querying a venue’s name and check who have visited this venue in a limit days. And the more work is to provide a hype link that can get more information about the user.

There are so many Issues I met when I was doing this problem; firstly, I need to look for a correctly API from both twitter and foursquare API documentations to get desire date. Secondly，I had to analysis data I got to retrieve the desire information(ie. Check-in Id),,after that, the largest issue is that some users cannot find through Twitter and foursquare API, so that it makes a large challenge during testing the system.

**Design and its Motivations:**

The design of this part use two APIs to get the correct output, first is using foursquare venue search API to get the venue’s coordinate and then use tweeter search API to get the users by querying the check in information.

**Requirements:**

This solution partly fulfilled the requirement. It can return a list of users who have visited the venue but cannot return the detail information of the user. limitations: In my opinion, the extension of this part is that the web may return a graph that shows the same venues these user both visited.!

**Limitations:**

## 2.2 Querying the Web of Data

**Issues:**

The aim of this section is attaching some additional function on the previous section. For the purpose of providing some information about other points of interest around the requested venues. We achieved this part by adding some hyperlink appended to the venue information of the last section. The problem we faced is that coping with the map demonstration and making the coordinates corresponding the location required.

**Design and its Motivations:**

**Requirements:**

**Limitations:**

## 2.3 Producing Data for the Web

**Issues:**

**Design and its Motivations:**

**Requirements:**

**Limitations:**

## 2.4 Storing information

**Issues:**

**Design and its Motivations:**

**Requirements:**

**Limitations:**

## 2.5 Web Interface

**Issues:**

For this part, we need to design an webpage for users to send their requests to the server. The page requires accessible, elegant and concise. Because the requirement of this program is divided into three individual part for each teammate, the biggest problem we faced is that associating different parts to the final webpage we designed.

**Design and its Motivations:**

**Requirements:**

**Limitations:**

## 2.6 Quality of the solution

## 2.7 Additional features

# 3 Conclusions