Bilkent University

Engineering Faculty

Department of Computer Engineering

CS 319

Object-Oriented Software Engineering

Course Instructor: Can Alkan

Semester: 2015-2016 Fall

Analysis Report

Group 04

Twitter Analysis

Miraç Aknar

Osman Koyuncu

Buğra Kalkan

Emir Özbek

28.10.2015

Contents

Table of Figures

1. Introduction (Problem Statement)

The system which takes information of Twitter users and stores them in a database, then analyzes those data, object oriented, and is “Twitter Analysis”.
This data analysis will be able to work by only taking user’s simple logging in to Twitter. After logging in the user will be able to choose different functionalities and get information about. After the process of taking, storing and analyzing data, the system will give its analysis result to user. This analysis is composed of followers’ and followings’ statistics, differentiated community analysis, like friend groups and families, finding popularity among user’s followers and followings. This system provides those by the analysis part as the control units of the system.

In this report, after introduction an overview of project will be given to reader with functional and non-functional requirements, constraints, scenarios and use case models and user interface, which is requirement analysis. After that as analysis, both object models, domain lexicon and class diagram, and dynamic models, state chart and sequence diagram is included. Finally, it ends with a conclusion.

2. Requirement Analysis

2.1. Overview

2.1. Functional Requirements

- The user should be able to login the program.
- When user login the program, user should use id and password.
- The user should be able to logout the program.
- The user should see their friend and their relation
- The user should learn their friends’ friends
- The system should determine the user friends
- The system should determine user friends’ friends
- The system should analyze some Twitter follow statistics
- The system should analyze Tweet statistics
- The system should analyze Twitter favorite’s statistics

- The system should analyze Twitter retweet statistics.
- The system should create Twitter community analysis
- User should see his/her own follow statistics.
- User should see his/her own Tweet statistics.
- User should see his/her own favorites statistics.
- User should see his/her own retweet statistics.
- User should see his/her own community analysis.
- User should see his/her follower's/ following user follow statistics.
- User should see his/her follower's/ following user Tweet statistics.
- User should see his/her follower's/ following user favorite's statistics.
- User should see his/her follower's/ following user retweet statistics.

2.2 Non-Functional Requirements

- The system shall be licensed.
- The system shall have firewall.
- There is "tutorial" document that explains how to use system to user.
- The system shall store maximum 1GB data.
- Data shall be imported into the system via LAN.

2.3 Constraints

- The program must be implemented in python.
- Database must be implemented in MySQL dB.
- The program must be online.
- Twitter api must be use (Tweepy)
- Graphic library

2.4 Scenarios

Scenario Name: Learning statistics about followers and followings

Ayça is a twitter user who posts tweets and spends time in twitter almost every day. He enjoys these activities but he wants to learn more about his activities such as at what rate he follows back his followers, what is the gender ratio of his followings, how many tweets does he post in a day on an average. He also wants to know these statistics of his followings or followers. He heard about a website in which he can find out these statistics. After an ordinary daily usage of Twitter, he connects to the website and see the welcome page. He logins with his twitter id and password. He finally sees the main page of the website.. When he clicks the button of follower's following's analysis. The page for analysis of his follower/following is displayed he can see many statistics that are mentioned above. Therefore he learned the information about his followers and followings that he wonders.

Scenario Name: Connection between social circles and users

Ahmet is a twitter user who received his high school education in İstanbul and comes to Ankara for his undergraduate education. He has many friends from İstanbul and he makes new friends in Ankara during his University education. So he has two different social circles and there are some connections between these circles. Twitter is a popular social media website among both of his high school friends and university friends. He wants to know about these circles and connections. He is also curious about who is more active in Twitter and whose posts are retweeted more or who follows who among his friends.

He finds out a website where he can see his social circles and connections between these circles even he can observe who is popular in this groups. He connects the website through his twitter account and clicked the button for his own analysis, he sees a graph where he can observe connections and circles.

2.5 Use Cases

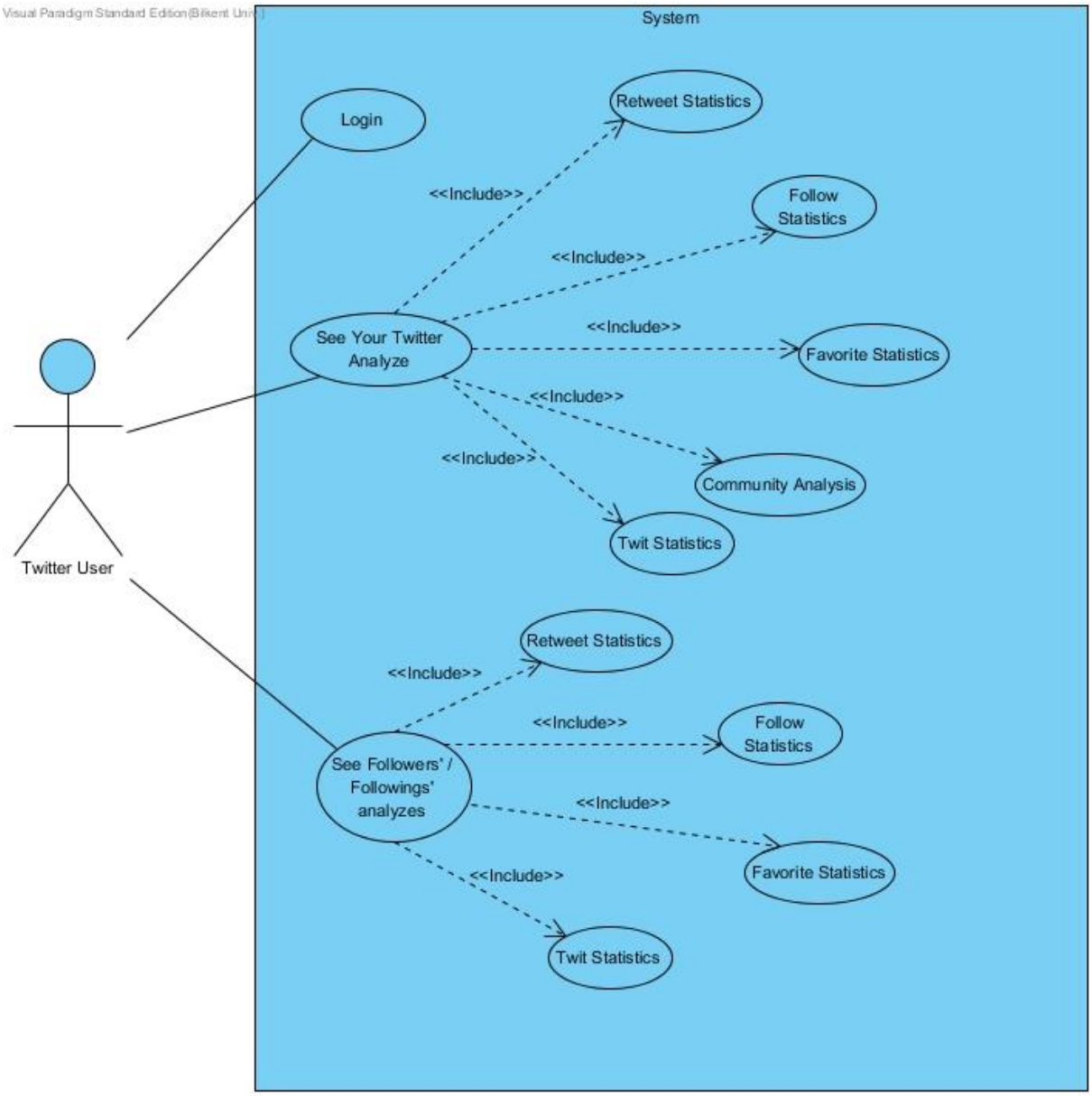

Figure1

2.5.1.Use Case: Login

Participating actor: User

Entry Condition: User access the website.

Exit Conditions: User exit the website.

Flow of Events:

1. User starts a browser.
2. User types the URL of the website on the address bar.
3. Welcome screen appears with a login panel.
4. User enters his/her twitter id and twitter password

5. User clicks the login button
6. User exit the website.

2.5.2.Use Case: Twitter Analysis for user

Participating actor: User

Entry Condition: User login to the system

Exit Conditions: User exit the website.

Flow of Events:

1. User login the system
2. User choose the button of Your Analysis
3. A page with five choices is seen.
4. User choose one of the analysis.
5. The analysis that use choose is screened
6. User exit the website

2.5.3. Use Case: Twitter Analysis for a follower or following of the user

Participating actor: User

Entry Condition: User login to the system

Exit Conditions: User exit the website.

Flow of Events:

1. User login the system.
2. User choose the button of a follower's following's analysis.
3. The search panel is seen.
4. User writes the name of a follower or a following.
5. User click the Analyze button
6. A page with four choices is displayed.
7. User choose one of the analysis.
8. The analysis that use choose is screened
9. User exit the website

2.6 User Interface

2.6.1 Login Screen

The initial screen of the program is the login screen. In this screen the user will enter his/her username and password to login. So the user is expected to enter correct username and password accordingly. After the user logins the second screen will appear.

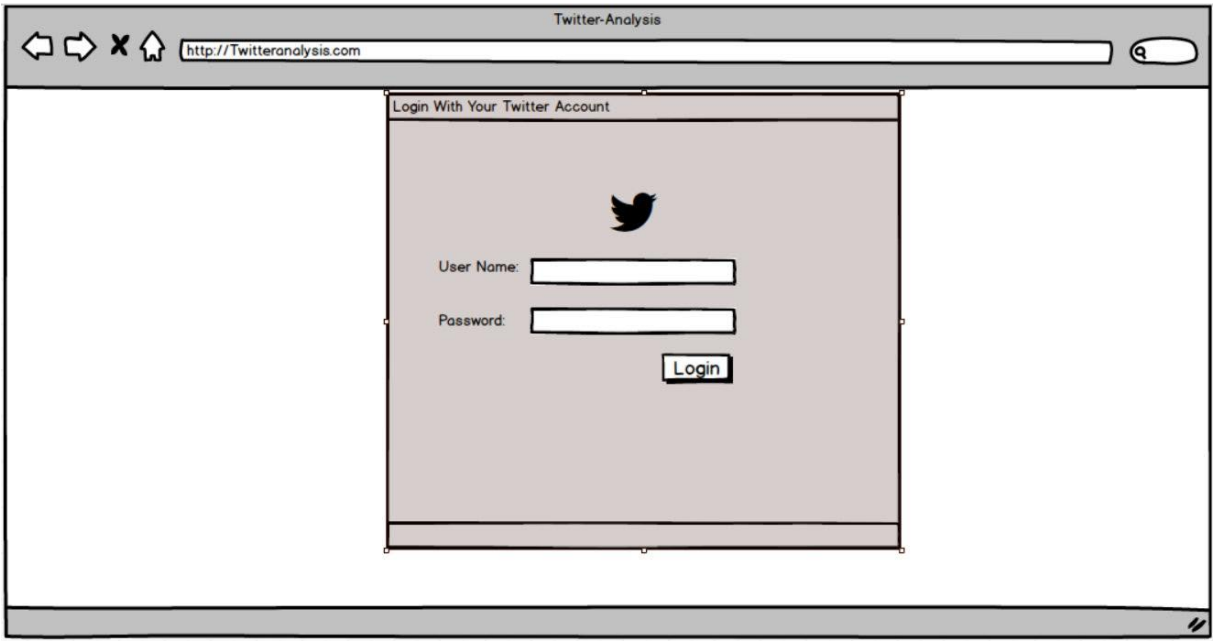

Figure2

2.6.2 Analysis Main Screen

This screen will appear right after getting login information of the user. The user will be able to get his own twitter analysis or one of his friends' analysis. If the user chooses 'Your Twitter Analysis' option, he/she will be asked to choose which analysis types are desired. If the user chooses the other option which is 'Follower's/Following's Analysis', he/she will be asked to pick which aspects of their friend's analysis are desired. The user also shall be able to logout safely.

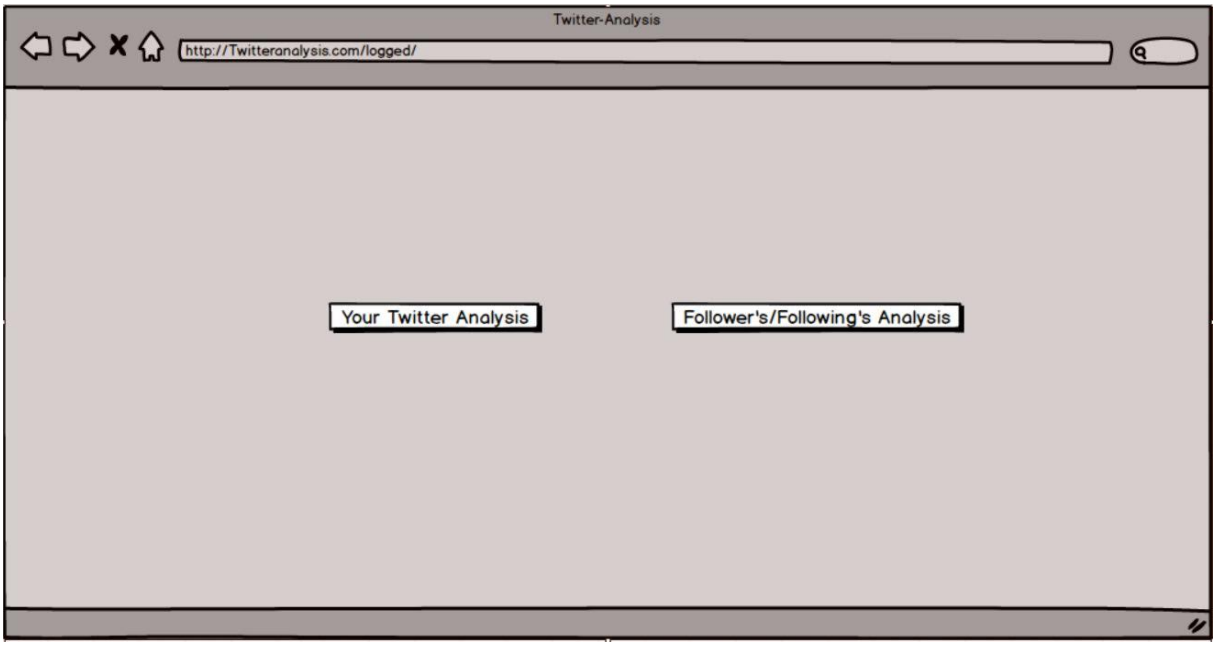

Figure3

2.6.3 Search Screen

This screen will appear if the user clicks on follower's/following's Analysis which means he wants to see one of his/her friend's analysis. In this screen, the user will be provided with a text label and a simple list. The user will be able to type one of his/her friend's name and see the analysis for that particular person. If the user types only name of that person, all the followings/followers who have the same first name will be shown in the list. If the user types the full name of the user, only that particular person will be shown on the list so that the user can easily reach his/her analysis.

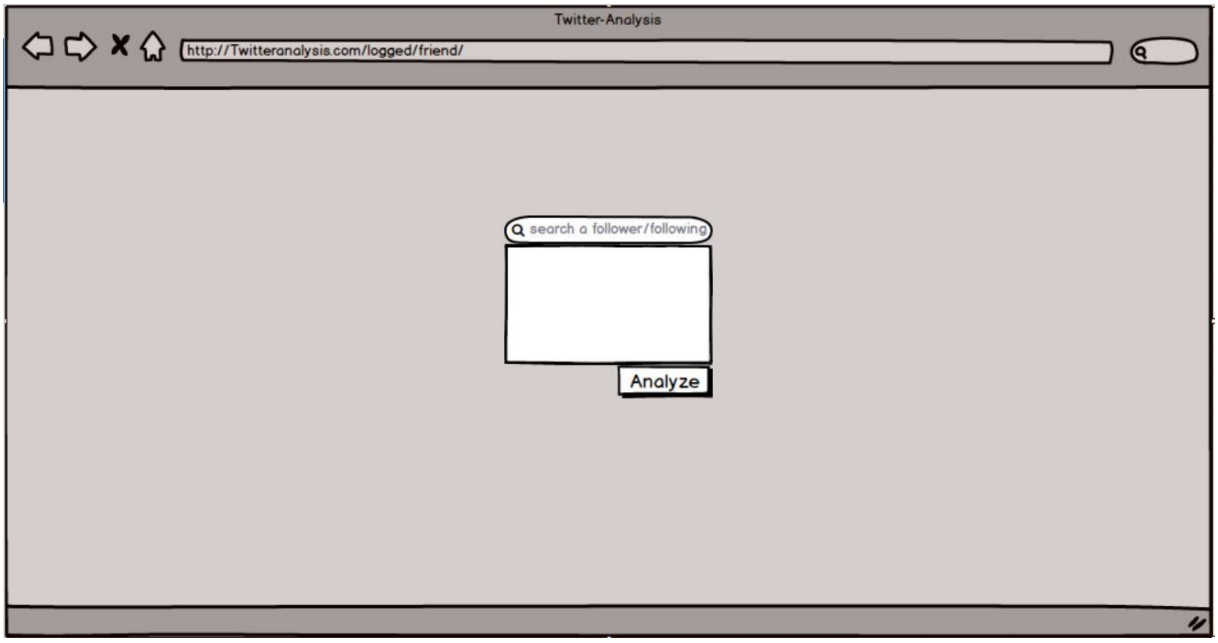

Figure4

2.6.4 Twitter Statistics Screen (Friend's)

After the user determines which of his/her friend's analysis he/she would like to see, this screen will be shown on the screen. At the left top corner of the current screen, the user will be able to observe his/her friend's profile picture, username, number of followings and followers accordingly. The user also will be given with 4 different options to continue with on their analysis process. If the user picks 'Follow Statistics' option, he/she will be able to see the screen which demonstrates the statistics of one of his/her friend's followings and followers. If the user picks 'Tweet Statistics', the Tweet Statistics screen will appear. If the user picks 'Favorite Statistics', the Favorite Statistics screen will appear. Lastly, if the user picks 'Retweet Statistics', the Retweet Statistics screen will appear.

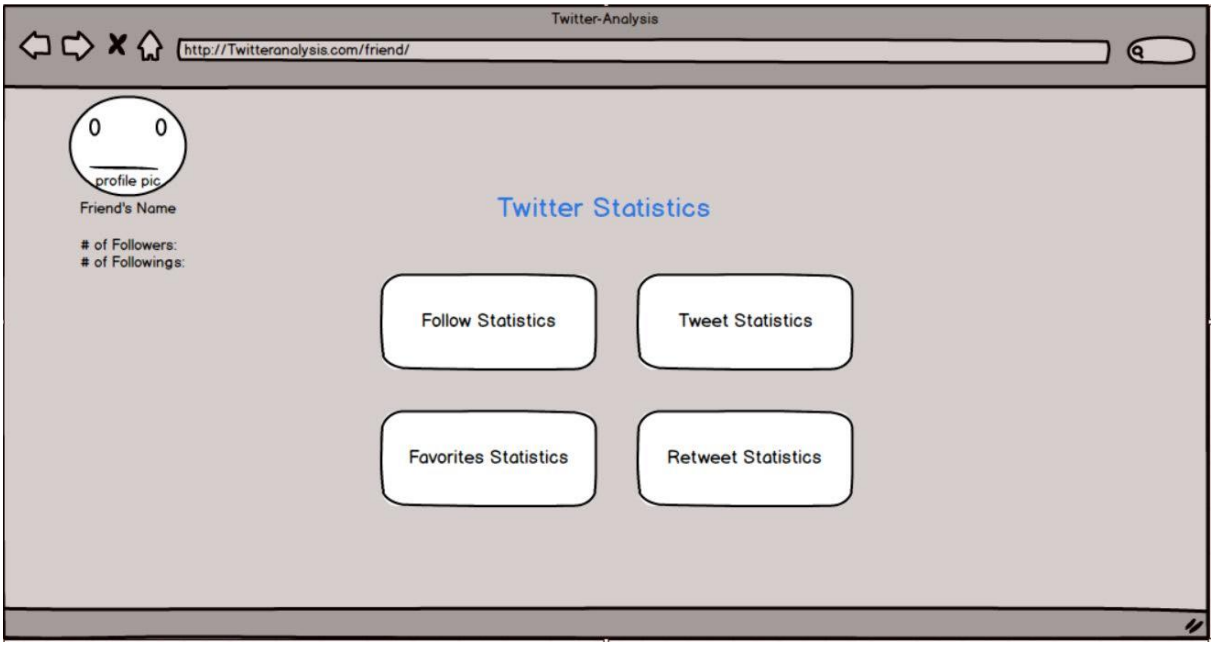

Figure5

2.6.5 Twitter Statistics Screen (User's)

If the user picks 'Your Twitter Analysis', this screen will appear on the screen. This screen has the same functionalities with the Twitter Statistics Screen(Friend's). But this screen also offers Community Analysis which is a very extensive type of analysis which analyzes following and followers and gives the information with a specific graph.

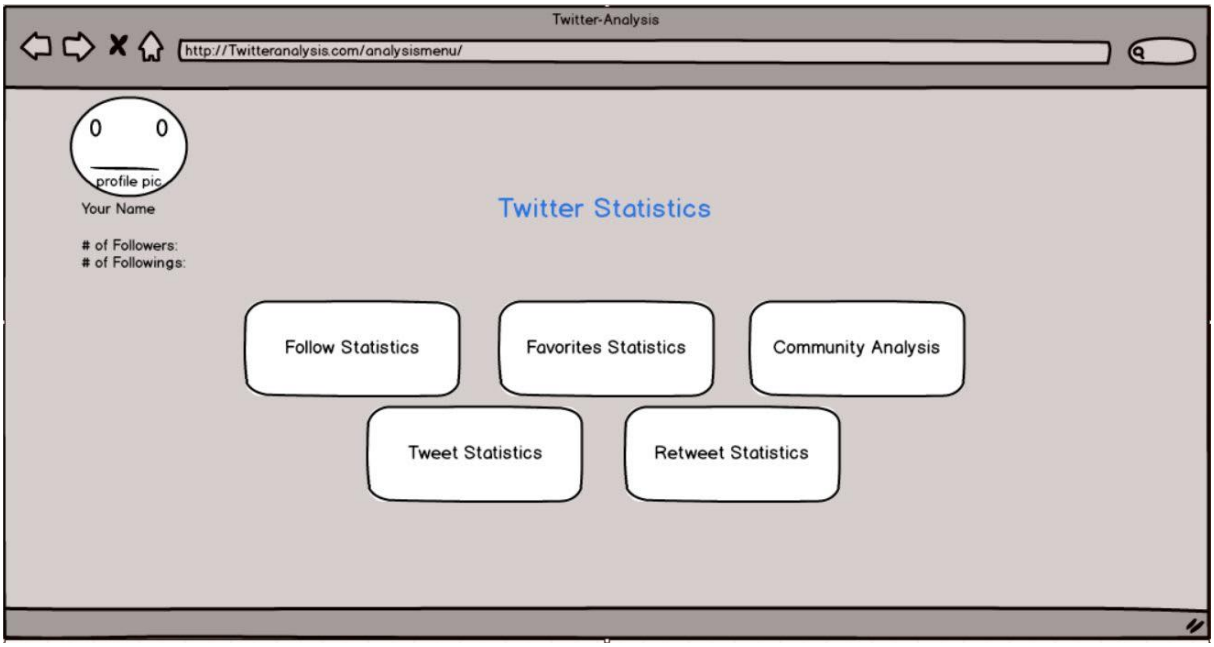

Figure6

2.6.6 Follow Statistics Screen

This screen allows the user to see both followings and followers statistics of a desired account. This screen has 4 diagrams (2 for followings and 2 for followers). In Followings section, the user will be able to see who follows and does not follow the desired account from his/her followings. The other graph shows how many of his followings are male and female. In Followers section, the user will be able to see who follows and does not follow the desired account from his/her followers. The other graph shows how many of his followers are male and female.

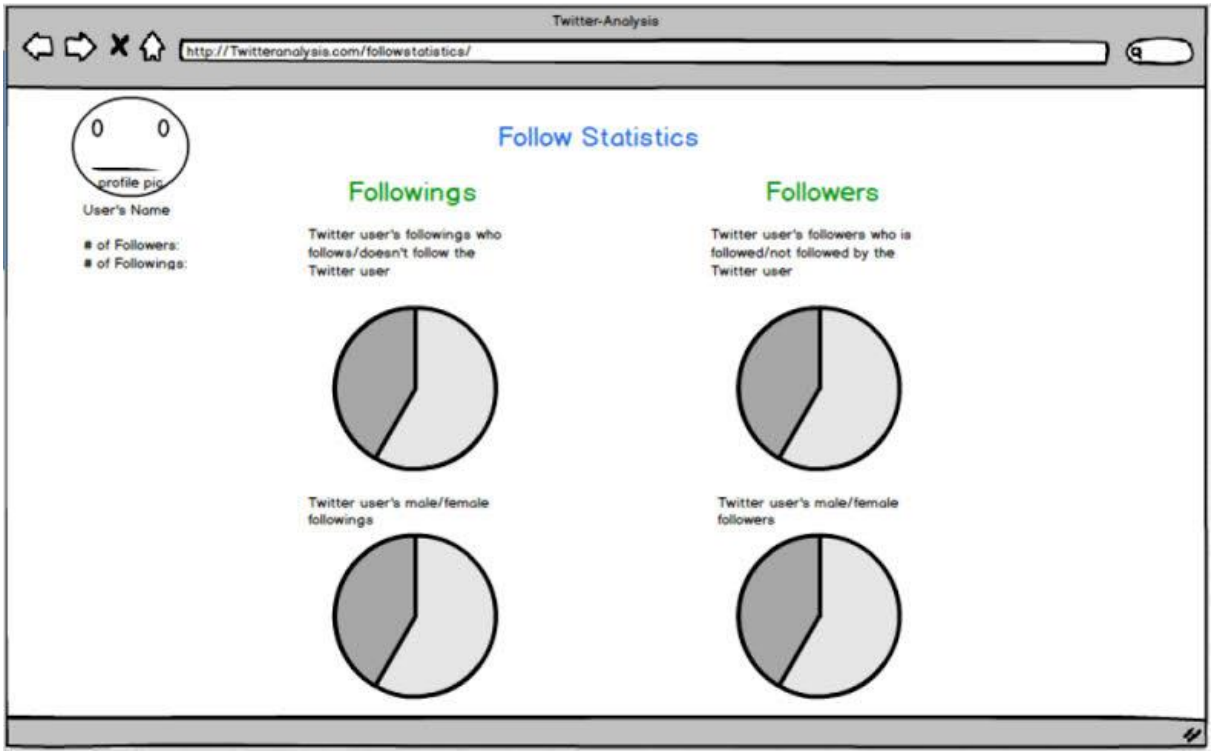

Figure7

2.6.7 Tweet Statistics Screen

This screen demonstrates two similar graphs to the user. According to last 20 posts of the account, the graph arranges tweets to days of weeks they are posted. The other graph also arranges last 20 posts of the account so that the user will be able to see which hours of day they are posted. Therefore, distribution of the last 20 tweets can be examined easily.

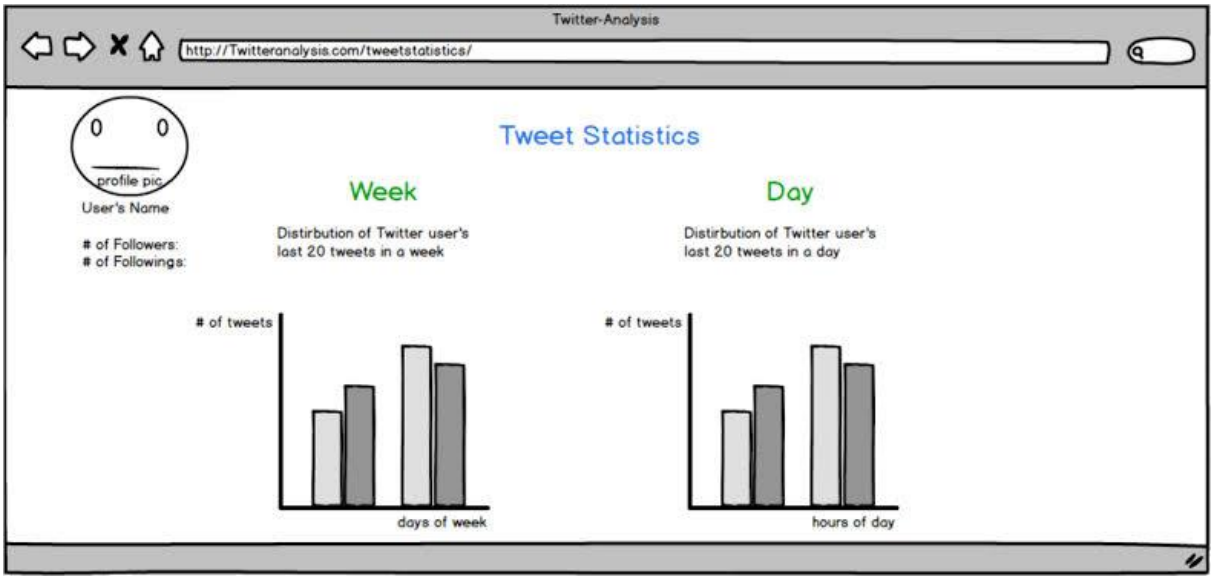

Figure8

2.6.8 Retweet Statistics Screen

This screen allows the user to see the last 10 tweets of the account and a graph which illustrates the frequency of followers who retweeted users' last 10 tweets so that the closest friends of the user can be seen easily. If a friend did not retweet one of the last 10 posts, he is included as never in graph. If a friend retweeted 1 or 2 tweets from last 10 post, he is included as rarely in graph. If a friend retweeted 3 or 4 tweets from last 10 post, he is included as rarely in graph. If a friend retweeted 5 or 6 tweets from last 10 post, he is included as often in graph. If a friend retweeted 7 or more tweets from last 10 post, he is included as frequently in graph.

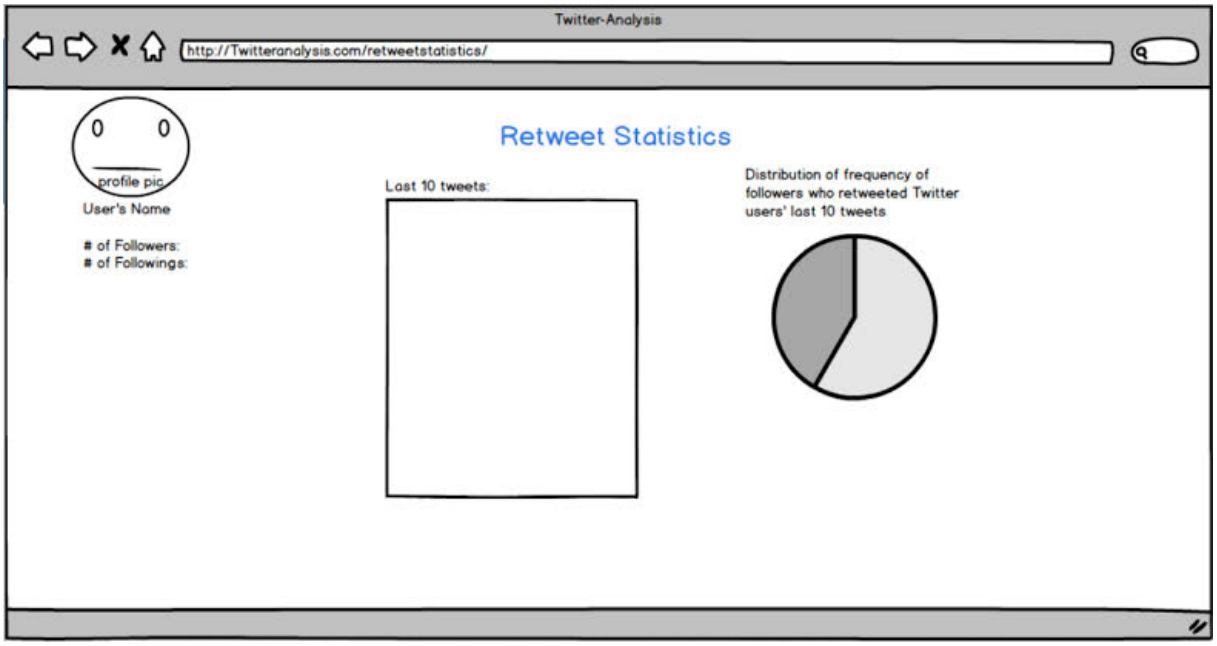

Figure9

2.6.9 Favorite Statistics Screen

This screen allows the user to see three of his friends who he/she favorited at most and a graph which illustrates the distribution of frequency of followings whose tweets the user favorited at most. This enables the user to figure out which of his/her friends he likes to favorite more than others.

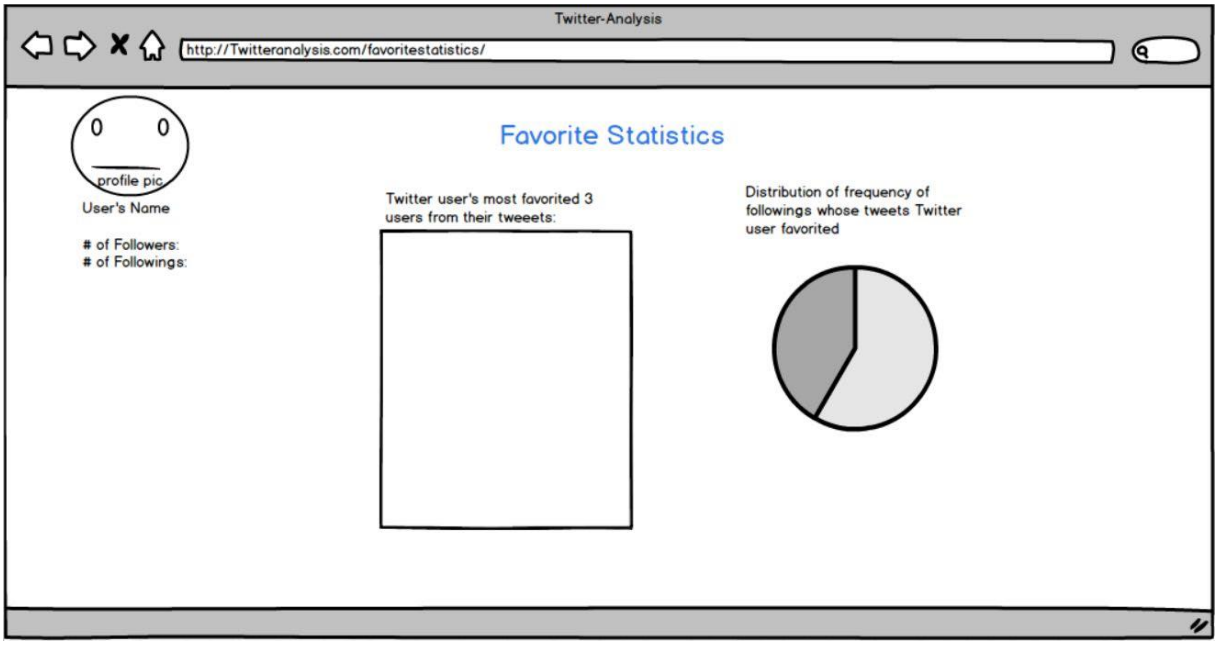

Figure10

2.6.10 Community Analysis Screen

This screen plots a graph which analyzes the user's friends in a way that close friend groups can be determined. Certain groups in his friends interact with each other more. Therefore they are treated as a social group in this graph. Interactions between friends are shown by lines. Friends are denoted by dots and their width and color determines how popular they are.

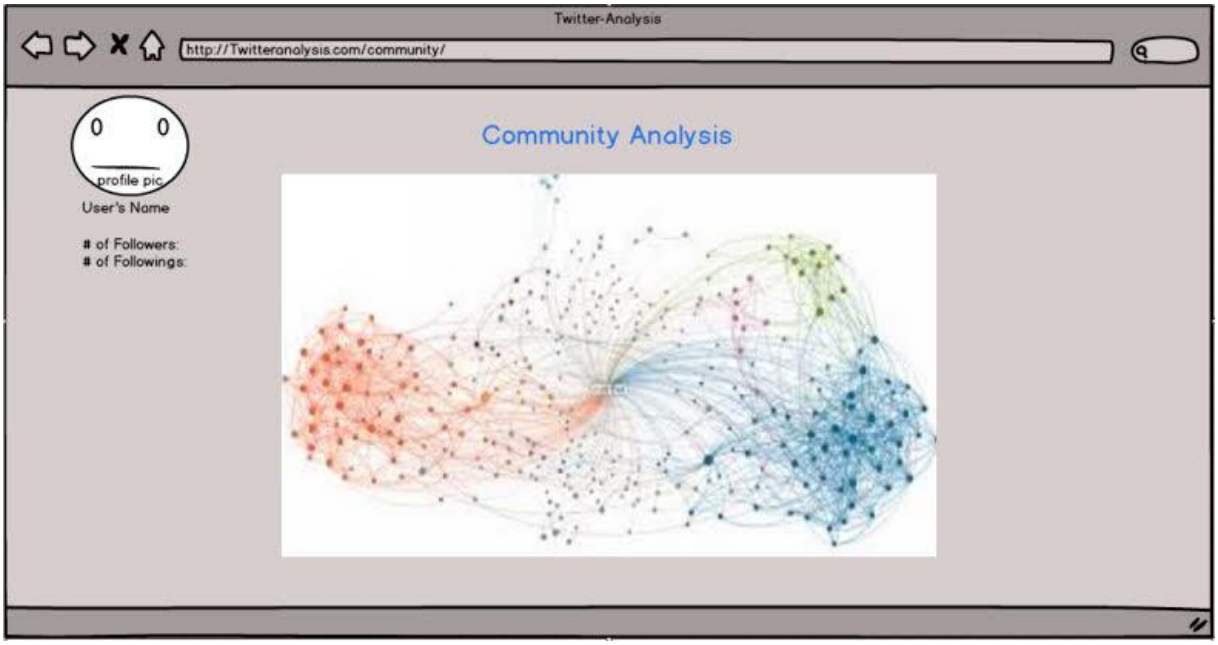

Figure11

3. Analysis
3.1 Object Model

3 .1.1 Domain Lexicon

The classes that are seen above are the model classes. There are 15 model classes.

Twitter User: The class that contains important features of twitter class which are twitter id, profile picture, followers and followings count and username.

Twitter Relation: It has source and target id. It represents a relation in which the user having the source id follows the user having the target id.

Tweet: This class contains a string containing the text and the id of that tweet.

Favorite: This class also has tweet, target id and source id which means the user having the target id like the tweet that belongs to target id.

Retweet: It has same attributes with favorite. Only difference is that one user retweeted another user's tweet.

Twitter User Group: It only has a list of twitter users.

Gender Ratio: It has an attribute named ratio, and one method which takes a group and returns a float that will be stored in ratio. It is the ratio of males/females.

Followers: This class is the child of the Twitter User Group. It is assigned to contain a group of users who are followers of a particular user and that particular user. It also have an instance of gender ratio.

Followings: This class has same attributes with Followers. The difference is it contains followings of a particular user.

File Object: It is the father class of the ones will be stored in file. Those classes are Twitter User Group, TwitterRelation and Tweet. The class has a dictionary has suitable elements to store in a file.

Database Object: It is the father class of the ones will be stored in database. Those classes are TwitterUser, Retweet, and Favorite. The class has a table name and list of strings to save in database properly.

Follow, Favorite, Tweet and Retweet Analysis: They have specified attributes which will be load by the particular controllers. The analysis results will be calculated according to data in the object of these classes

3.1.2 Class Diagrams

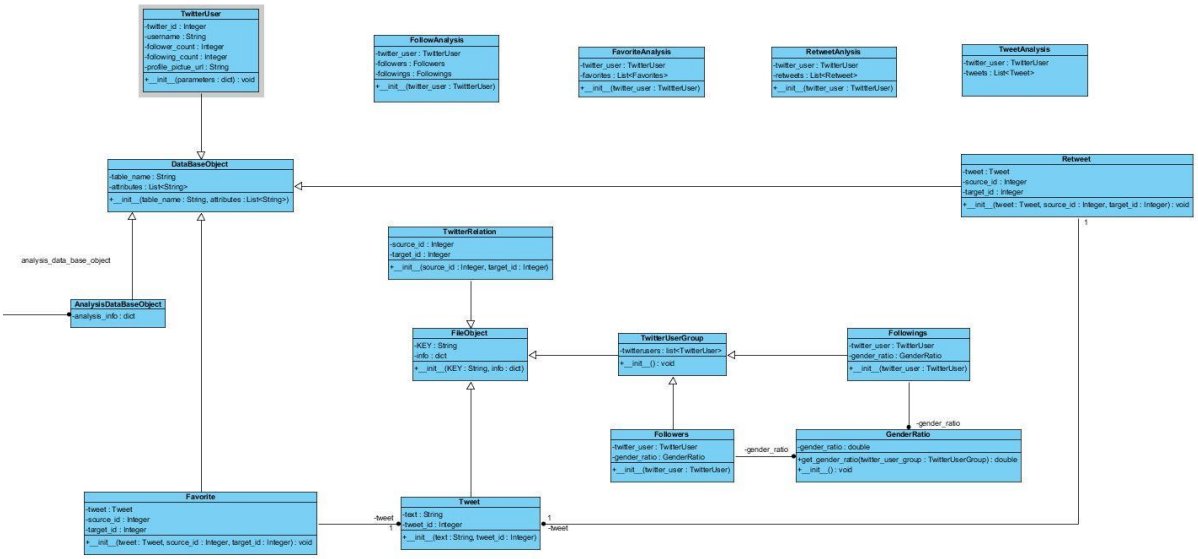

Figure12

.

The classes above are the model classes. The ones which are at the below are the control classes.

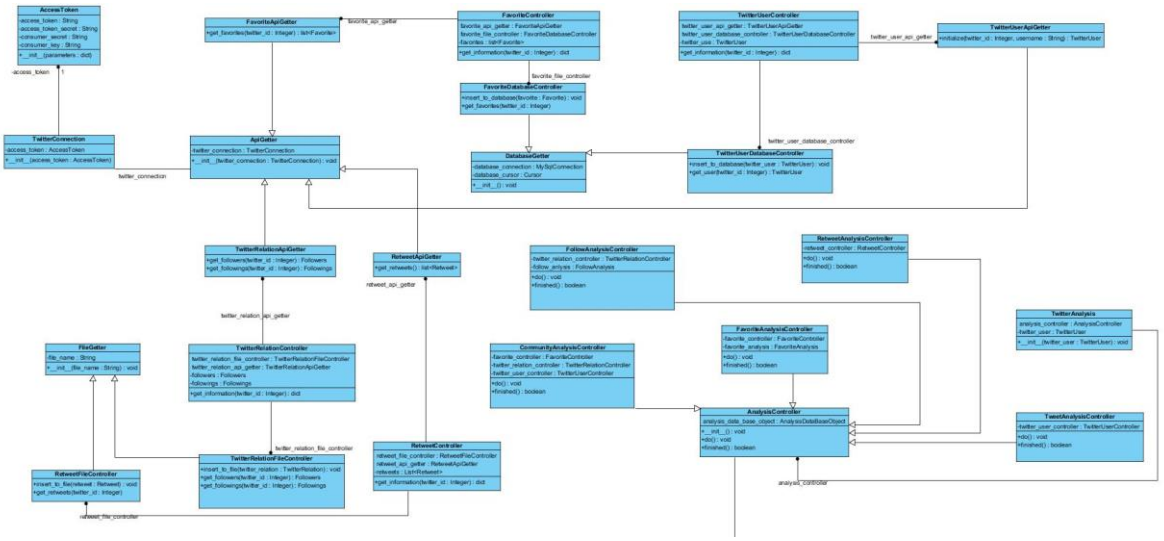

Figure13

3.2 Dynamic Model

3.2.1 State Chart

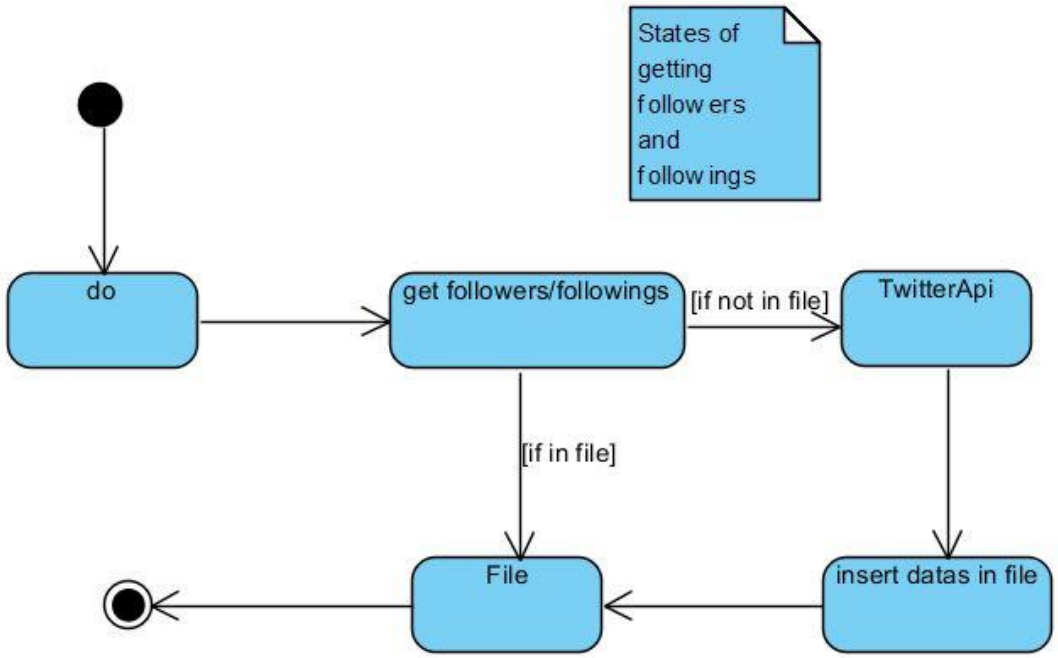

Figure14

This chart shows how to get followers/followings list from TwitterApi or file. do() is the function which starts all, then if the followers/followings wanted are already inside the file, finish, if not take it via TwitterApi then insert it in file.

3.2.2 Sequence Diagram

Scenario Name: New Analysis

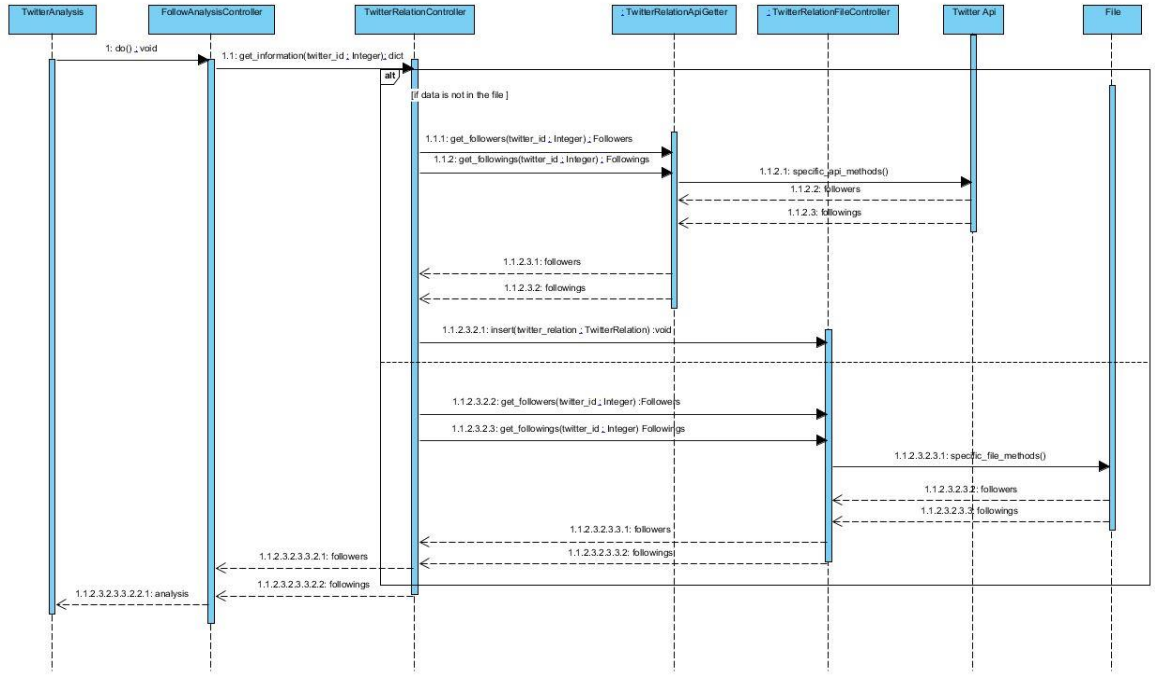

Figure15

The sequence diagram above is for follow analysis. For every analysis that user want, a Twitter Analysis object will be created. Then this object will have a controller according to the type of analysis, in this case Follow Analysis. For Follow Analysis, a Twitter Relation Controller will be created. Twitter Relation Controller will have Twitter Relation Api Getter and Twitter Relation File Controller.

If the data that we are looking for is not in the file, Twitter Relation Api Getter will take the data from the Twitter Api and take it back the relation controller. Then the data will be stored in the file. If the data is in the file at the beginning this process will not be happened. The data will be directly fetched from the file to twitter relation controller. Then it will be fetched into Twitter Analysis Controller. The analysis result will be evaluated in here and will be sent to Tweet Analysis.

The system will have three other analysis type: Tweet Analysis, Retweet Analysis and Favorite Analysis. These Analysis will differ only in storing. Tweet Analysis and Follow Analysis will

fetch the data from file, Retweet Analysis and Favorite Analysis will fetch the data from Database. The sequence diagram can be considered as that it belongs to other Analysis.

Conclusion

Twitter Analysis is an application which appeals to all kinds of twitter users. It provides both high level and in depth analysis including total number of followings, followers, retweets and favorites, all the circles and connections among a user's followers. There are plenty of existing similar applications, but this particular one offers different kind of analysis from others.

- User will be able to find the circles and connections from his friend's followers
- User will be able to store all the analysis
- User will have an opportunity to have a tutorial before using Twitter Analysis
- Graphics might be developed and new technologies might be added