
The Eurasia Proceedings of Educational & Social Sciences (EPESS), 2017

Volume 7, Pages 187-191

ICRES 2017: International Conference on Research in Education and Science

A DESIGN OF SOCIAL MEDIA ANALYSIS SYSTEM BASED ON MOBILE PLATFORMS

Aysegul Alaybeyoglu
Izmir Katip Celebi University, Computer Engineering Department

Levent Yavuz
Avrasya University, Computer Engineering Department

Abstract: Along with the developing technology, social media technologies have become widespread and the number of internet users has increased rapidly. In addition, social media platforms have become very popular and the number of active social media users has increased considerably. As a result of the increased use of social media, there has been a trend towards mobile platforms. In this paper, a design of a social media analysis system is developed using mobile platforms based on Android. By this way, important data about social media users will be able to gathered and analysed.

Keywords: Social media analytics, android, tweet analysis, user analysis

Introduction

With technology taking an important role in people's lives, some social media platforms such as Twitter have become popular. People can easily share their feelings and thoughts with other users using these platforms. These platforms provide the opportunity to interact with and influence other people. By this way, social media users can share simple text messages, pictures and videos via those social media platforms.

The Twitter platform is a social media platform with has most users from all of these social media platforms. Twitter is a social network where users can write "tweets", delimited by 140 characters and these tweets can be seen by other users too. Also, users can send not only text messages but also images videos, audio files and so on.

On this perspective, too much data is shared on Twitter, and when this data is analysed, meaningful results may emerge. Therefore, Twitter data can be thought that has important value to be analysed. Separately, social media platforms generally provide feedback on current events, companies, and other things and also provide some opportunities to evaluate their products or other things. When a new event occurs or a new technology emerges, many users start to tweet and retweeting about these events on Twitter. In this way, Twitter becomes a social data repository and information about the users and events can be accessed by this data repository.

The developed Android application accesses this data repository and performs data analysis. The analysis process is based on viewing and analysing all the data on Twitter about a keyword that the application user wants. With this mobile application, application users can view tweets on Twitter about any keyword and names of users' profile pictures, locations can be displayed as examples of users who posted tweet about this keyword. In addition to these, all tweets that are shared on Twitter in real time can also be displayed. These real-time tweets are displayed with user names and text content without any limitations.

The mobile Android application provides a system that collects, stores and analyses the tweets that users posted. This application is suitable for all Android-powered devices and requires internet connection so that the application can run without errors. When the application is started, it connects to the Twitter platform via internet and interacts with Twitter data. Due to the application is mobile based, it can be used easily wherever the internet connection is available regardless of the location. Today, with the development of mobile devices with

- This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

- Selection and peer-review under responsibility of the Organizing Committee of the conference

*Corresponding author: Aysegul Alaybeyoglu - E-mail: Email: aysegul.alaybeyoglu@ikc.edu.tr

powerful hardware resources, high level processing is possible on mobile devices. In this respect, similar mobile applications are needed and mobile application that performs social media analysis as an example.

Related Works

The emergence of mobile devices and the reduction of the cost of the internet made it easier for people to use more than one social networking service at the same time. Web search methods such as Google search are not enough for Twitter content search. Many companies, researchers, developers focused on tweets content analysis by applying some techniques on tweets.

In the business world, commercial companies, banks and similar organizations need systems that conduct social media analysis in order to receive feedback and improve themselves for the commercial or non-commercial products they have developed for potential customer. Governments also use similar systems to obtain information and feedback about on their own policies and on the opinions of their citizens. These example systems are generally server-based or desktop based applications.

Some works are being done and developed around Twitter analysis. In 2012 [1], a group of researchers identified a software architecture that aggregated tweets submitted in a given geographical location and within a specified time periods. This tweet gathering process was performed using Twitter Streaming API. In another study at Arizona State University [2], Twitter data was analysed and the detection of people using harmful substances was studied. The other study in 2012 [3], a group of researchers described their architecture to analyse Twitter data with using Twitter Rest API and Twitter Streaming API in PHP language and then MySQL database is used for storing data that coming from Twitter. In 2010 [4], a study is conducted to determine trends using the Twitter Streaming API. In this way, the most posted words in the tweets were instantly analysed on Twitter. In [5], tweets are posted by the users are analysed and displayed user's city level locations as results.

The common feature of all these studies is that they are implemented on desktop based platforms. These kinds of works require high speed internet connection and high quantity of hardware resource. Such studies are possible with cloud computing technologies [6] that are popular with the developing technologies. The cloud computing [7] technologies include inside of some technologies such as Apache Hadoop platform [8][9] Apache Spark platform [10][11]. As a result of the literature survey, there is no social media analysis work on mobile platforms. However, since the use of 4.5G has become widespread, mobile devices can perform as good quality work as some desktop devices.

Methods

Twitter provides API resources [12] to the usage for developers. Programmers are use the Twitter API to develop own applications or other projects that interact with Twitter. These applications talk with the Twitter API over Hyper Text Transfer Protocol and HTTP is a protocol that browser uses to visit and interact with web pages. For that reason, when using Twitter APIs, the internet connection is needed. Generally, these APIs are like as bridge between application and Twitter. There are many Twitter API libraries built for the Twitter platform. These API libraries have been implemented in different programming languages. Examples of these languages are Java, C++, ASP and C# programming languages. Within these libraries, Twitter4J library has been implemented in Java programming language.

Twitter basically provides two API to programmers. These are Streaming API and Rest API [13] [14]. The Streaming API is for application developers who want to receive a real-time stream of the public Tweets on Twitter. In this way, Streaming API provides to observe what is going on in the world. Tweet which is displayed with Streaming API is displayed in real time. The Rest API provides implementing access to read and write Twitter data. Rest API provides especially creating a new tweet, reading user profile and displaying follower data, and more. In addition to the Rest API, the Twitter API includes the Restful methods to send and receive twitter data. These methods are "Get" and "Post" methods. Get methods provides to obtain information such as user direct messages and post methods provides sending information such as sending to new direct messages from an application with Rest API which generally provides user information such as friendships, messages, statuses etc. so tweet text analysis is not suitable with Rest API. Both Streaming API and Rest API is important APIs for using and implementing Twitter data.

Twitter uses OAuth to provide authorized access to its API while sending secure authorized requests to the Twitter API. In this wise, Twitter provides security and standardization. Because, the users of Twitter are not

required to share their passwords and any other information with other applications or software. Many client libraries are compatible with Twitter’s OAuth implementation. There are two types Twitter API authentication model. These are user authentication and application-only authentication [15]. With application-only authentication, the applications can access Twitter server to obtain any information about it. But with user authentication, it is allowed to reach only user information such as reaching own Twitter profile, sending direct message, displaying coming message, displaying friends list and follower list etc. to making analysis on Twitter, application-only authentication is needed.

Twitter data is a repository about getting some answers from tweets written all around the world. When data is used correctly and effectively, Twitter data will carry a great value. To realize the analysis of Twitter data is needed to searching, storing, analysing and visualizing steps. Tweets are needed to be downloaded for an efficient data analysis. So that it uses Twitter API to collect tweets from Twitter server database. Analysing Twitter data focuses on two key aspects of Twitter data: network analysis and text analysis. Network analysis is based on actions between users about following, followers and retweets. It tries to find answers out of user network on Twitter and network analysis is important for individual analyses. Text analysis is based on tweets which users posted on Twitter. Text analysis tries to find tweet’s topics, any special words, owners and obtain any other information. For successful Twitter data analysis, searching, storing, analysing and visualization steps are required. After first three steps, Twitter data is downloaded, stored and analysed to obtain significant data. At this point, the purpose of analysis has completed. In visualization step, obtained Twitter data can be displayed via visualization tools.

The designed application is a mobile based android application that includes analysis and visualisation operations and uses the native API of Twitter. The mobile application gives services to a wide range of people, organizations or companies which need in public opinion pool about any subjects.

Design Of Social Media Analysis System Based On Mobile Platforms

The proposed system has four basic processes which are storing, searching, analysing and visualizing. Application searches Twitter cloud and results are downloaded from Twitter cloud. After that data goes into storing process to be stored in application’s data structures. In storing process, data may be applied to some filtering operations to eliminate spams, advertisements and other unrequired and unrelated data. Due to time limitations of Twitter API to reach its own database, crawling and storing operations are batch processing operations performed in background at regular intervals. The system architecture is shown at Figure 1.

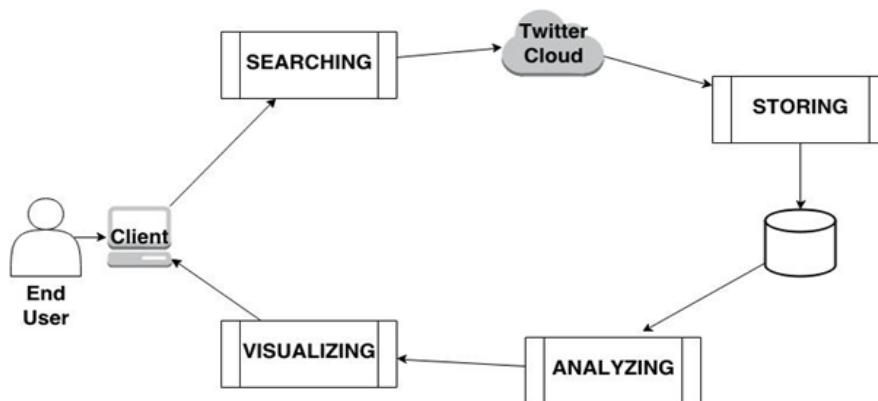


Figure 1. The proposed system architecture

The system collects data obtained from Twitter and analyses them in accordance with the requested topics. The proposed system:

- Allows users to collect Twitter data posted with given keyword.
- Allows users to collect real time Twitter data.
- Allows users to retrieve relevant tweets from the collected data using textual queries.
- Allows users to display user profile image.
- Allows users to search tweets and retrieve user information.
- Allows users to analyse tweets which posted on Twitter.
- Allows users to visualize the results of analysed tweets.

The use case diagram shown in Figure 2 illustrates how users are expected to interact with the system.

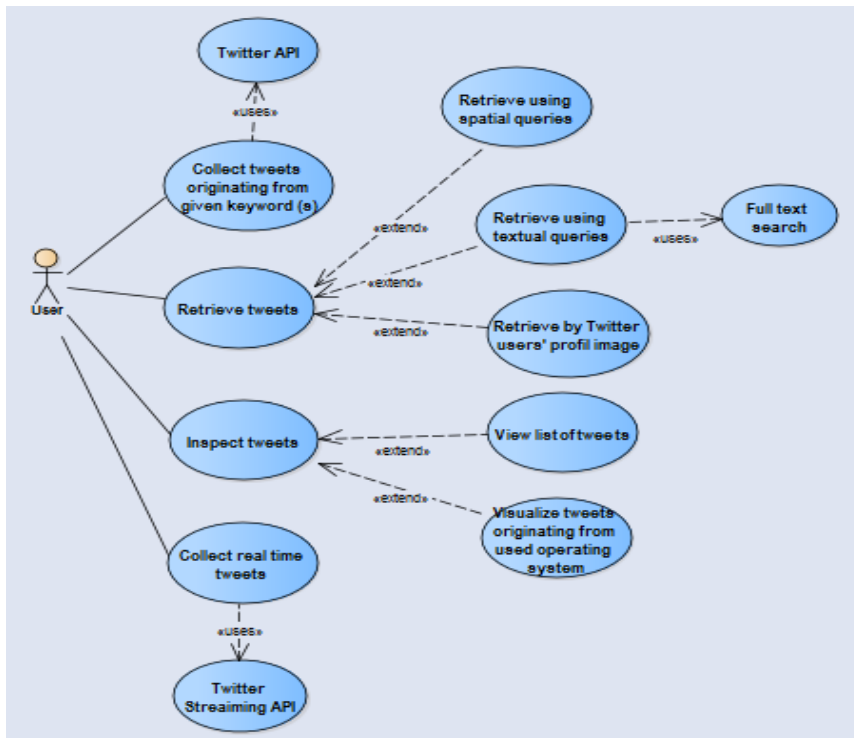


Figure 2. Use Case Diagram showing the expected usage of the proposed system architecture

The system has some functionalities. These functionalities are related with tweets obtained from Twitter. The first feature is getting tweet of a user. After a tweet is obtained, the same operation is continuously repeated, and the captured data is displayed on the screen as list view. Some pseudocodes of the system are given in Figure 3.

```

Get the tweets of a user:
Input Parameters: user id.
Returns: list of tweets(id)
Initialize tweetsIDs list with the tweets of the user input id
ListTweets
Sleep()
Foreach tweet in tweetsIDs
Begin:
Tweets.add(tweet id)
End:
Return Tweets.
End

Avoid collecting same data for same tweet:
Input Parameters: tweet id
Returns: Boolean, true if the tweet id exist, false otherwise
Begin:
List tweets;
Foreach user in users
Begin:
Tweets.add(user.tweetID);
End:
If (tweet id is in tweets)
Return true;
Else
Return false;
End
    
```

(a) Getting tweet of a user

(b) Avoiding to collect same data for same user

Figure 3. Some pseudocodes of the system

Conclusion

As a conclusion, technology developments in this century made communication and blogging shifts through social websites and platforms like Twitter. And this made social media platforms such as Twitter a valuable information resources in the important events like natural disasters, elections.

On this paper, it was aimed to gather important information via Twitter by making a smart crawler. This crawler will collect and store tweets and related information based on search with given keywords and other parameters. Then, the collected tweets are analysed with textual measures and a simple graphic chart was provided to the end

user as visualization. With the application design, the data coming from Twitter is planned to be evaluated in meaningful way. This application is hoped to give the important clues in about many things such as the people's social behaviour, personalities and characters. At this point, a lot of data can be obtained which may have beneficial results on behalf of our country. As the years are progressed, Twitter and similar social platforms will become even more popular and the number of active users will increase. So, the data obtained from social platforms will be very important and as the technology continues to evolve and social media usage increases, the more of this kind of work will be required.

References

- Oussalah M., Bhat F., Challis K., Schnier T., (2013), A software architecture for Twitter collection, search and geolocation services, *Know.-Based Syst.* 37, 205-120
- Tian Q., LAgissetty J., Li B., (2016), Finding Needles of Interested Tweets in the Haystack of Twitter Network, *IEEE International Conference on Advances in Social Networks Analysis and Mining*
- Black A., Mascaro C., Gallagher M., Goggins S., (2012), Twitter Zombie: Architecture for Capturing, Socially Transforming and Analyzing the Twittersphere
- Sakaki T.i Okazaki M., Matsuo Y., (2010), Earthquake shakes Twitter users: real-time event detection by social sensors, In *Proceedings of the 19th international conference on World Wide Web (WWW '10)*
- Cheng Z., Cavarlee J., Lee K., (2010), You are where you tweet: a content based approach to geo-locating twitter users, In *preceedings of thw 19th ACM international conference on Information and knowledge management*, pp. 759-768
- Demchenko Y., Bernstein D., Belloum A. Oprescu A., Wlodarczyk W., Laat C., (2013), New Instructional Models for Binding Effective Cirricula on Cloud Computing Technologies and Engineering, *IEEE International Conference on Cloud Computing Technology and Science*
- Bisdikian C., Mitschang B., Pedreschi D., Tseng V., (2011), Challenges for Mobile Data Management in the Era of Cloud and Social Computing, *IEEE International Conference on Mobile Data Management*
- Singh T., Darshan V. S., (2015), A Modern Data Architecture with Apache Hadoop, *IEEE International Conference on Green Computing and Internet of Things*
- Manikandan S. G., Ravi S., (2014), Big Data Analysis using Apache Hadoop, *IEEE International Conference on IT Convergence and Security*
- Backhoff O., Ntoutsi E., (2016), Scalable Online-Offline Stream Clustering in Apache Spark, *IEEE 16th International Conference on Data Mining Workshop*
- Maheshwar C. R., Haritha D., (2016), Survey on High Performance Analytics of Bigdata with Apache Spark, *IEEE International Conference on Advanced Communication Control and Computing Technologies*
- Twitter Application Page, (2017), <https://apps.twitter.com>
- Twitter Streaming API, (2017), <https://dev.twitter.com/streaming/overview>
- Twitter Rest API, (2017), <https://dev.twitter.com/rest/public>
- Twitter OAuth, (2017), <https://dev.twitter.com/oauth>