

# EXPLOITING TWITTER IN MARKET RESEARCH FOR UNIVERSITY DEGREE COURSES

Zhenar Shaho Faeq<sup>1</sup>, Kayhan Ghafoor<sup>2</sup>, Bawar Abdalla<sup>3</sup> and Omar Al-rassam<sup>4</sup>

<sup>1</sup>Department of Software Engineering, Koya University, Koya, Iraq

<sup>2</sup>Department of Software Engineering, Salahaddin University, Erbil, Iraq

<sup>3</sup>Department of Software Engineering, Koya University, Koya, Iraq

<sup>4</sup>Department of Mathematics, Koya University, Koya, Iraq

## ABSTRACT

*Twitter is one of the biggest microblogging platforms that millions of Tweets of maximum size 140 characters are posted every day by its users. Recently many research efforts have been presented towards analyzing the content of information in Twitter for marketing purposes. However, existing analysis approaches have shortcomings in terms of exploiting this media to improve marketing policy in a University. Thus, efficient analyzing technique is a pressing need to complement the existing approaches and hence improve the performance of marketing department in a university through getting public mood from Twitter and apply a number of analysis. To this end, in this paper, we develop a desktop based application that can help marketing staff in Heriot-Watt University to analyse Tweets related to the university and can help in course marketing. The application enables users to retrieve Tweets from Twitter corpus which are related to the university. Afterwards, users will be able to perform a number analysis on the retrieved Tweets such as Word Frequency Analysis, User Frequency Analysis and Sentiment Analysis. Finally, users of the application will be able to export the results in to a spreadsheet. After the application is evaluated, most of the participants agreed that it is reliable and easy to use and that the application has a good performance and can be used for marketing purposes.*

## KEYWORDS

*Microblogging, Marketing, Word Frequency Analysis, User Frequency Analysis, Sentiment Analysis.*

## 1. INTRODUCTION

Twitter was first built in October 2006 and has grown very rapidly which by 2012 reached more than 200 million users [1]. Nowadays, Twitter has the maximum number of users among the other microblogging services. Therefore it has become a new platform for most businesses and marketers to do marketing in an interactive and dialogic way and share information with their consumers [2].

Furthermore, having a huge number of users in Twitter and rapid growth of this number on a daily basis and the fact that most of tweets and user profiles are public, these reasons have drawn researcher's interest to conduct more researches and analyse Twitter feeds [3][4][5]. Twitter feeds has been analysed for many purposes and in various fields of science including politics, business, human health, news media, marketing etc. The numbers of research carried out on Twitter are increasing gradually due to the importance of this platform on having massive amount of data from various geographic places.

In general, Twitter plays a major role in marketing development that many companies have taken benefits of this medium for enhancing marketing policies. For instance, according to Guglielmo (2009) cited in [6] Dell Inc. produced \$6.5 million in profits by using Twitter accounts in 2009 in orders for accessories, personal computers and software. Another example is Zappos which is a clothing and fashion company uses Twitter regularly in order to generate and manage customer relationship through having more than 1,200 conversations per month with their customers. Moreover, a company like Tiny Prints has increased its position in search ranking by 47% through using Twitter [7].

Most approaches face difficulties due to the limited Tweet feed which is provided freely by Twitter. The Twitter APIs offer limited amount of Tweet feeds so when a user attends to retrieve data repeatedly, he/she will encounter rate limit problems. Moreover, lack of having a concrete algorithm for analyzing the sentiment of Tweets in c# leads to having accuracy problem while applying sentiment analysis. In order to overcome these problems, a new approach is used to combine two Twitter APIs that can retrieve maximum amount of recent and real time Tweets available freely by Twitter company. In addition, two lists of positive and negative opinion words or sentiment words for English are used in order to perform an accurate sentiment analysis. Positive list contains 2006 positive words and negative list contains 4783 negative words. Therefore, sentiment of the Tweets will be determined based on the existence of these words in the Tweets. The aim of the project was to develop a desktop based application that can help marketing staff in Heriot-Watt University to analyse Tweets related to the university and can help in course marketing.

The application is developed by c# programming language. In this application, Twitter search API and Streaming API are combined together so that maximum available Tweets related to a search keyword will be retrieved to the application. The retrieved Tweets will be displayed in a chronological order from oldest to newest. The paper is organized as following: chapter two of the report contains a literature review on how twitter has been used and analysed as a tool for marketing purposes. In the third section, the approach and design required for the project is identified. Section five of this paper describes the details of implementing various analysis. Finally, fifth section contains a conclusion for the project.

## **2.ROLE OF TWITTER IN MARKETING**

Increasing use of Twitter and participation of users from all around the world have increased number of brands to have an active account in Twitter and connect with their consumer. According to [1] there are more than 100 million active users that Tweet frequently and there are 250 million Tweets per day. Therefore, brands have a great opportunity to find and communicate with their consumers and share new information. According to [8] there are two main factors that business value and use Twitter as a communication platform for marketing, having a huge audience size and user attention.

Since Twitter is a conversational platform, it provides the opportunity for both brands and consumers to have a dialogic environment for exchanging information between them which is very useful for both side. Dickey and Lewis (2009) cited in [6] believes that using Twitter for marketing is relatively faster, less expensive and geographically unbounded as compared to other marketing platforms. Moreover, [9] predict that 73% of companies that use Twitter for marketing purpose will increase their activities.

According to [2] there are four main purposes of contributing Twitter as a marketing tool:

- Research in marketing and generating feedback for brands
- Branding , publicity, and reputation management
- Creating network with businesses
- Provide customer service and managing customer relationship

Therefore, Twitter platform can be used by any kind of companies or organisations including universities. For instance, Heriot-Watt University can take the advantage of Twitter in various ways; firstly for spreading information about new courses, events, university's current position, seminars etc. secondly, Twitter can be used for establishing a good relationship with students and public. Thirdly, Heriot-Watt University can use twitter to get general sentiment of public concerning the university in general and particular courses. Finally, this platform can be used of advertising purposes.

However, on the difficulty of identifying the contents of the Tweets and choosing those which are related to marketing, [10] conducted a research to choose Tweets that involve possible data related to any company and categorize the Tweets into two group depending on whether they refer to a company or they describe a different subject which is not related to any company. Finally it was difficult to be accurate and efficient in distinguishing between the Tweets which were referring to a company and the ones which were not. [10] indicated that complexity of identification and grouping of Tweets is due to some reasons:

- Tweets are formed from short text (maximum 140 characters) which is more difficult as compared with long texts.
- Most of the users tend to write using informal writing style.
- Using poor grammatical structure and unfamiliar vocabularies.

Therefore, analysing and grouping of Tweets cannot be always very accurate; however, they are still very efficient methods that help in making new decisions giving directions.

The usefulness of Twitter as a marketing tool might differ among the companies depending on size and popularity of the company, since big companies naturally will have more followers as compared to smaller ones. [11] conducted a research in order to examine whether Twitter can be used for determining marketing strategies of Small and Medium-sized Enterprises (SMEs). This research carried out in order to clarify three main points

- How Twitter is used by SMEs for marketing strategies
- Investigate the importance and restriction of Twitter as a marketing tool for SMEs
- Future plan of the SMEs in using Twitter for improving marketing strategies.

The researchers used a qualitative-exploratory methodology and conducted over a sample of five Romanian and British SMEs. The findings indicated that Twitter is one of the most useful tools that can be used and relied on by companies to contact and listen to their consumers in order to share information and affect their idea. However, companies should not rely completely on Twitter and use it as the only communication tool with consumers.

In order to assess the role of Twitter in building relationship between companies and consumers, [12] performed a study on business accounts in Twitter in order to check whether these businesses use Twitter for interacting and communicating with their consumer. The study analysed 1000 business Tweets which were chosen from 5.245 business accounts. The results indicated that

business Tweets were mainly used for two purposes, a large portion of Tweets was posted for spreading news updates about business (mostly through providing a link), the other large portion were used for interacting and engaging with users individually, which was usually in a form of conversation to provide customer services. In addition, a portion of the Tweets of the companies were used for announcement including new offers or deals, events and job opportunities.

In order for the companies to have a continuous good relationship with their consumers or followers, every company need to specify some individuals or a group in order to continuously interact with them. [13] conducted a research in order to investigate how Twitter is used by corporations in term of communication and a tool for building relationships. The study analysed the Twitter profiles of 47 companies and worked on creating a guideline to improve practitioner's knowledge about using Twitter more effectively for communication. In addition, the study concentrates on the use of Twitter by people who practice public relations in order to create effective and long-term relationship. After a content analysis which has been done on the Tweets of the corporations, the data indicated that companies were using Twitter in order to communicate with publics. Moreover, the two most significant features which indicate having a relationship between companies and consumers were trust and control mutuality. The Tweets showed that the companies want to create a mutual relationship with consumers through providing a friendly dialogic atmosphere.

Stock markets are among the other markets which take advantages of Twitter platform for developing their marketing strategies. [14] carried out a study in order to predict the behaviour of some stock market indicators through analysing posts from Twitter. The study analysed a huge collection of random Tweets of a six month period, it measured hope as a positive and fear as a negative mood and compared it with the stock market indicators such as Dow Jones, VIX, S&P 500, and NASDAQ to find the relation in between. Finally, there were a negative correlation between the Tweets and other stock market indicators such as Dow Jones and NASDAQ; however, the research indicated that most of the Tweets had a positive correlation with VIX. Therefore, the stock markets can be predicted by investigating the emotions of users in Twitter. From the researches, it can be stated that Twitter can be used for marketing in three main ways. Firstly it can be used to enhance brand-consumer relationship as of changing information and ideas between them. Secondly it can be used for marketing new products by brands. Finally it is a very important tool in a way that people's sentiment can be analysed about a particular product, subject or behaviour.

Therefore, marketing staff in Heriot-Watt University are able to enhance relationships with students and public through providing a conversational media by using their Twitter account. This can be achieved by answering questions, giving directions, providing required information etc. In addition, these accounts can be used for marketing and advertising available courses in the university and drawing public's attention toward the courses.

### **3. PROPOSED METHODOLOGY FOR DATA COLLECTION FROM TWITTER**

Concerning our application, it is working based on two RESTful Twitter APIs, Twitter Search API which belongs to REST API and public streams in Streaming API. The main reason for choosing these two APIs in developing our application and not the other available APIs is that both of them allows user to search for their required keyword(s) in public Tweets. They provide Track parameter so that users will be able to search for specific Tweets. In addition, the two APIs retrieve Tweets from public Twitter corpus rather than focusing on individual accounts. The concept of working of these APIs which will be used in developing the application is the same which is retrieving public tweets. However, they differ from each other based on the publishing date of tweets.

When a user enters a keyword(s) in the application, the keyword(s) will be used as the track parameter for both Twitter search API and Streaming API. Best practice is to enter maximum 10 keywords since Twitter Search API accepts limited number of keywords. However, developers can separate the keyword source for the Streaming API, for instance, declaring constant keywords internally in the codes or providing a separate textbox for entering keywords for Streaming API since capability of track parameter of Streaming API is better than Search API as it accepts up to 400 keywords.

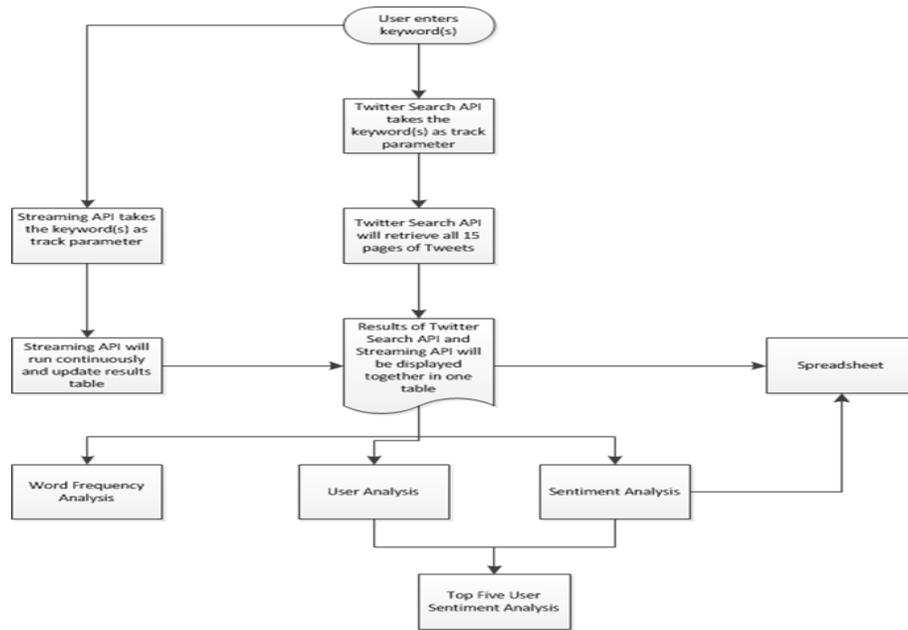


Figure.1. Proposed approach for developing the application

The result table of the application will be updated every time Streaming API fetches a new Tweet. As it was mentioned earlier, several analysis will be performed based on the results table, such as word frequency analysis, user analysis and sentiment analysis. In addition, in order to perform top five user sentiment analysis, sentiment and user analysis needs to be performed in advance. Moreover, the results of both APIs can be exported to a spread sheet file before performing sentiment analysis and after. If results are saved after performing sentiment analysis the spread sheet file will contain sentiment of authors as well.

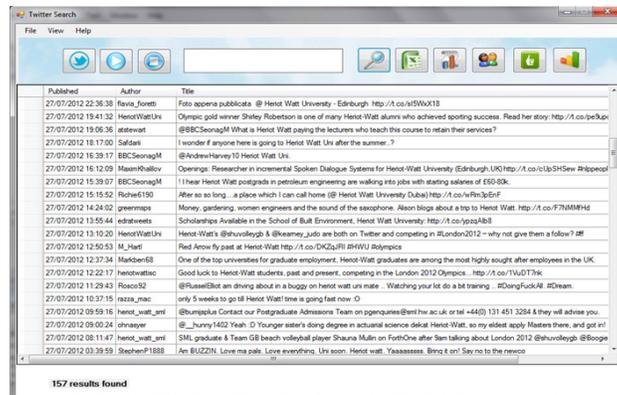


Figure 2: Results of a search function

### 3.1.Integration of Both Search and streaming APIs

In order to get a vast amount of Twitter feeds which will consist of recent and up to date Tweets, it is recommend for developers to use both Twitter search and Streaming APIs while developing an application. Therefore in developing our application, both Twitter search API and streaming API are integrated to get the maximum number of Tweets possible.

Twitter search API is one of the earliest API of Twitter which is specialised for carrying out searches concerning recent tweets. It provides an index containing recent tweets rather than complete index of all tweets. Therefore, the index of tweets provided by Twitter search API contains tweets of last 6-9 days. Developers cannot use This API to find Tweets older than 9 days [7].

Twitter streaming APIs are among the most powerful APIs that Twitter provides for developers. Streaming APIs are a set of APIs that helps in providing global stream of Tweet data for developers. Twitter streaming APIs retrieves real-time Tweets rather than recent Tweets; this makes Streaming APIs different from Twitter search API [15]

## 4.AUTHORS ANALYSIS ON THE RETRIEVED RESULTS

### 4.1.Abstract Word Frequency Analysis

One of the main functionalities provided in the application is word frequency analysis. This function counts the frequency of the words repeated in the retrieved Tweets and it will give a general idea to users on what are the most repeated words by public. Therefore, by applying this functionality to the retrieved results, users will get an overview on what people are mentioning and discussing the most. After performing this function, users will see a table consist of the words and their frequency of repeating. In addition, users can find location of any word in the retrieved Tweets by just double clicking the required word. When a user double clicks on any word in the wordlist, the Tweet(s) will be highlighted that contains selected word. This might help marketing staff for identifying the purpose of usage of a particular word.

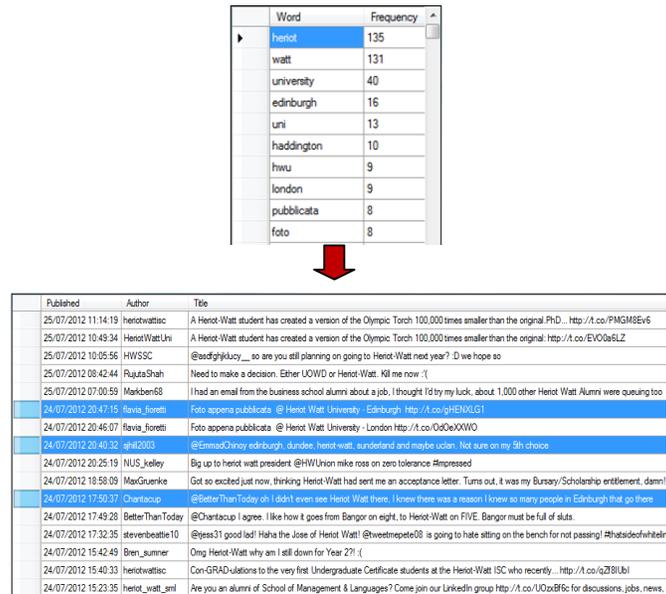


Figure 3: Display the frequency of repeating words and the usage of specific words in the contents of the Tweets

before displaying wordlist results in a table, a stopwords list is used to filter the words in the wordlist which does not help in processing of natural language, for instance short function words, such as on, is, at, the etc. and grammatical words such as am, I, are, she etc. For this purpose, a default English stopwords list is used which covers most of unwanted words [16].

### 4.2. User Frequency Analysis

User frequency analysis helps marketing staff and user perform some analysis on the users of the retrieved Tweets. The function will list author names in a descending manner based on number of tweets an author has tweeted. Through using this functionality, user will be able to check who are the authors that are Tweeting the most about a specific keyword such as HWU. The result of this function will be a table containing author names and their frequency of repeating in the results table.

In addition, users are able to highlight Tweet(s) of a particular author by just double clicking on the author name in the user frequency table. This will help users of the application to understand how a particular author thinks about a specific keyword(s).

### 4.3. Sentiment Analysis

Sentiment analysis is one of the main functionalities of this application which will be used to classify tweets based on sentiment of people which may be positive, negative or neutral. Users can determine how public thinks about a specific keyword such as HWU. In order to apply this functionality to the application, two lists of positive and negative opinion words or sentiment words for English are used. Positive list contains 2006 positive words and negative list contains 4783 negative words [17]. Therefore, when a user clicks on sentiment analysis button, every Tweet in the result table will be checked whether it contains a word from positive or negative word lists. If a Tweet contains a positive word exist in the positive word list, the background color of the Tweet will change to Green, if a Tweet contains a negative word exist in the negative word list, the background color of the Tweet will change to red. Finally if a Tweet does not contain any word from the positive and negative word lists, its background color will remain in white. However, determining sentiment analysis of the Tweets only using these two word lists is not sufficient for an accurate sentiment analysis. It is recommended to have multiple tools working together to determine sentiment of Tweets in order to achieve more accurate results

Published	Author	Title
28/07/2012 13:06:30	PadonHouse	I posted 3 photos on Facebook in the album "Padon House & Heriot-Watt Uni Collaboration 2011/12" <a href="http://t.co/G5OovvqD">http://t.co/G5OovvqD</a>
28/07/2012 12:54:34	LewisM11	@fergusath Heriot Watt U training with anyone
28/07/2012 11:39:00	tafashondary	Good Luck to the Heriot Watt University Students that are competing in the @London2012 ! <a href="http://t.co/Pq2L504">http://t.co/Pq2L504</a>
28/07/2012 05:58:25	FarySanabli	@kuhenza ohh so you'll be having fun with cousins eh? And btw I heard Heriot watt is another IES? :o is that true?
27/07/2012 22:36:35	flavia_jureti	Foto appena pubblicata @ Heriot Watt University - Edinburgh <a href="http://t.co/7EZZ15W8">http://t.co/7EZZ15W8</a>
27/07/2012 22:36:38	flavia_jureti	Foto appena pubblicata @ Heriot Watt University - Edinburgh <a href="http://t.co/s5WAX18">http://t.co/s5WAX18</a>
27/07/2012 19:41:32	HeriotWattUni	Olympic gold winner Shaley Robertson is one of many Heriot-Watt alumni who achieved sporting success. Read her story <a href="http://t.co/pe3up">http://t.co/pe3up</a>
27/07/2012 19:06:36	astewart	@BBCSeonagM What is Heriot Watt paying the lecturers who teach this course to retain their services?
27/07/2012 18:17:00	Safdar1	I wonder if anyone here is going to Heriot Watt Uni after the summer.?
27/07/2012 16:39:17	BBCSeonagM	@AndrewHarvey10 Heriot Watt Uni.
27/07/2012 16:12:09	MaximKhalov	Openings: Researcher in incremental Spoken Dialogue Systems for Heriot-Watt University (Edinburgh, UK) <a href="http://t.co/c1Up5H5ev">http://t.co/c1Up5H5ev</a> #hippeo
27/07/2012 15:38:07	BBCSeonagM	!! I hear Heriot Watt postgrads in petroleum engineering are walking into jobs with starting salaries of £50-80k.
27/07/2012 15:15:52	Richie6190	After so long... a place which I can call home (@ Heriot Watt University Dubai) <a href="http://t.co/wRm3pEnF">http://t.co/wRm3pEnF</a>
27/07/2012 14:24:02	greenmaps	Money, gardening, women engineers and the sound of the saxophone. Alison blogs about a trip to Heriot Watt. <a href="http://t.co/F7NMMH4d">http://t.co/F7NMMH4d</a>
27/07/2012 13:55:44	edratwosta	Scholarships Available in the School of Built Environment, Heriot Watt University. <a href="http://t.co/ypza4b8">http://t.co/ypza4b8</a>
27/07/2012 13:10:20	HeriotWattUni	Heriot-Watt's @shuvolleygb & @keamey_judo are both on Twitter and competing in #London2012 - why not give them a follow? #F
27/07/2012 12:50:53	M_Hart	Red Arrow fly past at Heriot Watt <a href="http://t.co/DKZuRl">http://t.co/DKZuRl</a> #HWU #olympics
27/07/2012 12:37:34	Markben68	One of the top universities for graduate employment, Heriot-Watt graduates are among the most highly sought after employees in the UK.
27/07/2012 12:22:17	helenwattinc	Good luck to Heriot Watt students, past and present, competing in the London 2012 Olympics. <a href="http://t.co/71uDT7K">http://t.co/71uDT7K</a>
27/07/2012 11:29:43	Roseco92	@RussellRiot am drivers about in a buzzy on heriot watt uni mate... Watchers your lot do a bit training... #DontFUCKU #Dream

Figure 6: Analyzing sentiment of users

#### 4.4. Top Five User's Sentiment Analysis

After performing sentiment and user frequency analysis in the application, users will be able to see a chart that contains sentiment of top five users that have Tweeted the most in the retrieved results. This function can be applied after making sentiment and user frequency analysis. The chart will visualize sentiment of top five users in user frequency analysis.

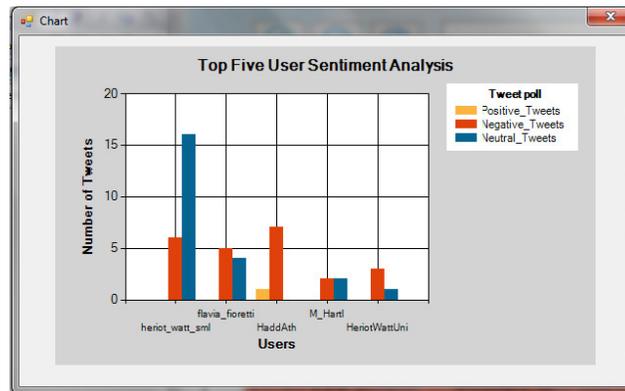


Figure 7: Chart illustrating top five users' sentiment

#### 4.5. Exporting the Results into spreadsheet

Users of this application will be able to export the results obtained by the application to a spreadsheet for further analysis. Moreover, results can be exported before and after applying sentiment analysis, if results are saved after performing sentiment analysis the spreadsheet file will contain sentiment of authors as well.

### 5. CONCLUSION

The application provides some functionality such as retrieving the Tweets related to the university and enable users to make several analysis such as sentiment analysis, user frequency analysis, word frequency analysis etc. Thus, in order to listen to a vast amount of Tweets related to Heriot-Watt University and getting public mood concerning the university in general or some particular department, this application can be used.

The application which is introduced in this Paper enables HWU marketing staff to retrieve recent (6-9 days old) and real-time Tweets from Twitter data based on keyword(s) that will be specified by the staff while starting the search process. Afterwards, the results obtained by the application can be analysed by some functions provided in the application. For instance, user frequency analysis can be used to determine the users who are Tweeting the most about HWU or any other particular keyword and this function will help marketing staff check how a particular user thinks about their specified keyword in the search. Moreover, sentiment analysis can be used to check the sentiment of the users and the results can be displayed in a chart showing the sentiment of top five users that have Tweeted the most about the specified keyword(s).

In conclusion, the developed application can be used by marketing staff in HWU In order to improve the marketing policy of the university through getting public mood and information concerning the university from Twitter.

## REFERENCES

- [1] Twitter1, (2012) What is Twitter? Also available at: <https://business.twitter.com/en/basics/what-is-twitter/> (last Accessed 20/02/2012)
- [2] Thoring, A. (2011) 'Corporate Tweeting: Analysing the Use of Twitter as a Marketing Tool by UK Trade Publishers', *Publishing Research Quarterly*, 27, 2, pp. 141-158.
- [3] Poblete, B. et al (2011). Do all birds tweet the same? : characterizing twitter around the world. Proceedings of the 20th ACM international conference on Information and knowledge management. Glasgow, Scotland, UK: ACM. p:1025-1030.
- [4] Ghafoor, K. Z.,et al (2011). A Fuzzy Logic Approach to Beaconing for Vehicular Ad hoc Networks. *International Journal of Telecommunication Systems*. 53(4), 330 – 342. SpringerLink.
- [5] Khokhar, R. H. et al (2011). Fuzzy Assisted Social-based Routing for Urban Vehicular Environments. *EURASIP Journal on Wireless Communications and Networking*. Hindawi Publishing Corporation. Vol. (178), No. 1.
- [6] Kwon, E.S. Sung, Y( 2011) 'FOLLOW ME! GLOBAL MARKETERS' TWITTER USE', *Journal Of Interactive Advertising*, 12, 1, pp. 4-16, Business Source Premier, EBSCOhost, viewed 11 July 2012.
- [7] Twitter2, (2012) Using the Twitter Search API. Also available at: <https://dev.twitter.com/docs/using-search> (last Accessed 03/07/2012)
- [8] Laudon, K.C. Traver, C.G (2010). *E-commerce 2010: business, technology, society*. New Jersey: Prentice Hall.
- [9] Stelzner, M. A. (2011) "Social Media Marketing Industry Report: How Marketers are Using Social Media to Grow Their Businesses," [socialmediaexaminer.com, http://www.socialmediaexaminer.com/social-media-marketing-industry-report-2011/](http://www.socialmediaexaminer.com/social-media-marketing-industry-report-2011/).
- [10] Pérez, F. et al (2011) On the Difficulty of Clustering Microblog Texts for Online Reputation Management. In: *Proc. ACL-HLT 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, WASSA-2011*, June 24, Portland, Oregon, USA.pp 146-152
- [11] Bulearca, M. Bulearca, S (2010) 'Twitter: a Viable Marketing Tool for SMEs?', *Global Business & Management Research*, 2, 4, p. 296.
- [12] Popescu, A. Jain, A ( 2011) 'Understanding the functions of business accounts on Twitter', *International World Wide Web Conference*, p. 107
- [13] Edman, H. (2010) *Twittering to the top: A content analysis of corporate tweets to measure organization-public relationships* (Master's thesis). Louisiana State University
- [14] Zhang, X. et al ( 2011), 'Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear"', *Procedia - Social And Behavioral Sciences*, 26, The 2nd Collaborative Innovation Networks Conference - COINs2010, pp. 55-62
- [15] Roden, T (2010) *Building The Realtime User Experience: Creating Immersive And Interactive Websites*. Beijing ; Cambridge : O'Reilly
- [16] Stopwords.(2012)EnglishStopwords. Also available at: <http://www.ranks.nl/resources/stopwords.html> (Last Accessed 015/07/2012)
- [17] Hu, M. Liu, B (2004). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. Seattle, WA, USA, ACM: 168-177.