

Twitter Data Streaming and capturing for Tourism Dashboard application

H.S. Albusaidi^{1*}, P.K. Udupi², V. Dattana³

¹Department of PG Studies, Middle East College, Muscat, Oman

²Department of Computing, Middle East College, Muscat, Oman

³Department of Computing, Middle East College, Muscat, Oman

*Corresponding Author: hsbusaidi15@gmail.com

Available online at: www.isroset.org

Received 8th Nov 2016, Revised 28th Nov 2016, Accepted 20th Dec 2016, Online 30th Apr 2017

Abstract— Information and communication technology intervention, advancement in internet technology narrowed down the geographical barrier for information Interchange [1]. Tourism sector is one of the primary benefitted sector due to rapid changes of information technology. Introduction of smart tourism, online transactions, visually stunning virtual world creation of tourism places, availability of immediate information further accelerated the expansion of tourism sector and driven tourism sector from traditional environment to technology driven business environment. Social media usages, tourism dashboards, media streaming ensured to achieve greater success for tourism sectors [2]. This paper identifies and evaluates methods of information capturing and processing pertain to tourism sector. The paper also explores the possibility of social media data streaming, extraction and mining [3] with the future scope of research towards visualization of information through dashboards.

Keywords: Twitter API, Twitter sentiments analysis, Data analytic, Dash Board, Smart Tourism, Big data, Oman Tourism, Oman economic, Twitter, Twitter API, Twitter hashtag.

I. Introduction

Many of the world's nations seeking to introduce modern technologies for the development of the tourism sector. More than 10% is the total contribution of travel and tourism of whole global economy GDP [4].

This vital sector is the source of income for many countries in the world such as Spain, Malaysia, France and other countries, where it is estimated that generated from tourism to these countries by more than \$ 72 billion in 2014, [4] which makes this sector strong export income to revive the economies of these countries.

Spanish government has sought to exploit the Big data in the development of the tourism sector, as well as the other such as Dubai in the United Arab Emirates, where tourists can now be defined, hotel occupancy and knows how congestion these cities through streaming data across systems used in them.



Figure 1. The World of Tourism

In light of this great momentum of the data streaming, it easy for tour companies put tourist programs via the Internet, which contains a lot of tourists to search for tourist destinations, search for places to spend their holidays. TripAdvisor is one of the tourism business portal, where more than 50 million visitors visits this site per month to search for hotels, restaurants, flights occasion and other things that can be done by sitting at one place and also there are also many other social networking sites and internet sites, which caters the need of tourists.

II. Review of existing system

There are many studies done, that talked largely about the importance of the Sentiment of tourists when they visit the tourist destination around the world. Few study also focused on the role of tourists' feedback in the enhancement of religious tourism destination. One of the study done by researchers also focused in how to benefit from the analysis of the feelings of tourists for their visit to 'Sacred Mount of Oropa' and 'Sacred Mount of Orta' (Piedmont Region, Italy) and how it can be used to manage to improve performance in the tourism services available in this destination.

The study has focused on the Sentiment of the tourism destination visitors through analysis of TripAdvisor website (www.tripadvisor.com) through the analysis of the reviews that posts by website visitors.

With regard to the first case study, the 226 reviews were posted on TA by people who visited the Sacred Mount of Oropa. The data were collected on 25th November 2014 and processed in December 2014. For the second case study, the 243 reviews posted by visitors to the Sacred Mount of Orta were collected on 10th April 2015 and processed during May 2015. [4]



Figure 2. Existing System Process

It's so flexible to extracting data from websites such as TripAdvisor which provide a huge amount of data and visitors reviews, but there are many defects of the methodology that used in this system where accompany this process is the most important that this system depend on the analysis of a specific amount of data does not exceed 500 samples while can extract all of the data

in more than a similar site in the Italian Oropa. This existing system is using a manual methodology to analysing data, which uses a conventional analysis tools for that, and that which is common, but they do not provide a live data.

Present day, analysis methodologies may differ where they can extract a live data which provided by the social networking like twitter and Facebook and other sites.

III. Proposed System for Twitter data analysis for tourism

The huge amount of data streaming daily through smart devices and applications running on this devices, such as social media and websites have made the business sector is connected permanently with customers, [6] therefore it has become very easy to know the customer's sentiment then analysis it to improve the level of services. [7]

All touristic countries, have dealt with the data generated by social media to derive the views of tourists for every tourist area, and where that Twitter is one of the most active social networking sites [8]. Many data extraction groups consider it as an important informational wealth, which this data can mining then analysing to identify the behaviour of tourists.

Use Twitter API Streaming to extracting data through Twitter is a new revolution in the data mining, where can easily extract it or stored in databases, and then can be used in the multiple purposes, including future decisions in the tourism sector making. [9]

The tourism sector is one of the more growing of business sectors in the world [10], so it is important to know the Sentiment of tourists at all tourist places in the country, so many countries around of world are Applying the analyse the Sentiment of tourists and find out how satisfied they are when they visit tourist reality.

Twitter is one of the wealthiest social networking with data and the most popular in the world wide web. With an estimated number of users are more than 50 million of users, with more than 8 terabytes of data generated per day and more than 70 million tweets are posts every day and 800 tweets per second. The hashtag make it easy to look to the hot trend at the time, and easily extracted and analysed.

The research Twitter data streaming and capturing for tourism dashboard focuses on the analysis of the sentiment of tourists in the Sultanate of Oman through the extraction of data from Twitter and stored in databases and analysed through the proportion of containment tweets on the positive and negative words and displayed on Dashboard. The result analysis will allow the decision-makers in Omani tourism sector to take appropriate decisions to improve the quality of tourism services in Oman.

3.1 Data Mining

Twitter is one of the most important social networking sites which is active and richer live data that generated by users per day, and as the hashtag collects all tweets under the label and one therefore extracted, it's simple if used the appropriate tools. In this research, all data will be extracted and analysed from #hashtag directly in real time using twitter API streaming at the time when the hash is active

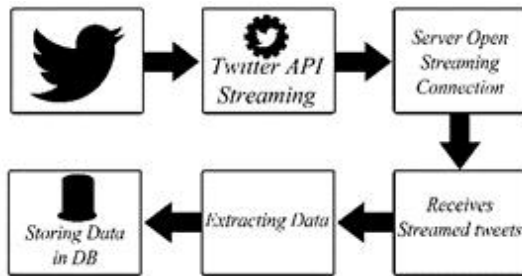


Figure 3. Twitter data extraction

3.2 Data Storing

After extracting the data from the Twitter #hashtag are stored in SQL Server database, SQL Server is a product provided by a Microsoft to store and manage data, typically a "relational database management system" (RDMS). The out then displayed in the dashboard after the entering key words that represent the positive and negative tweets. The visual dashboard displays done after filtering this data and represented as graphs in dashboard to show the satisfaction about the tourism services.

3.3 Data Analyzing by using key words (Positive and Negative)

During the process of data extraction, mining and storing and using specific applications, the data refined by entering the keywords that contains the negative words and positive words which used by twitter users. The application automatically retrieves tweets from twitter hashtag and assembling it in the database, then sorting tweets between tweets containing positive and negative words and then displayed in the dashboard as a graphs which represents negative and positive tweets.

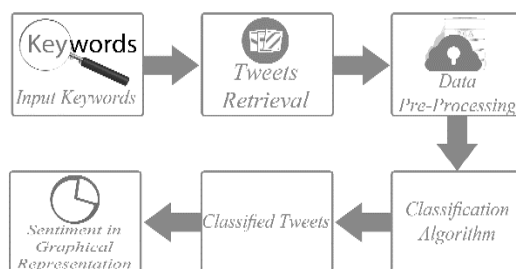


Figure 4. Dashboard Visualization

3.4 Development Dashboard

The dashboard is a tool for enterprise data management, which displays live data to help decision-makers to build appropriate decisions.

There are two types of dashboard, analytical and operational. Analytical dashboard used to display statistics as a visible data, the second type of dashboard, is the Operational dashboard that display alerts and processes. Dashboard application can be developed using Asp.Net with programming language C# or any other suitable software tools. This application connected with SQL server database for data retrievals.

IV. Conclusion

The outcomes of this paper is focused on measuring the satisfaction of tourists of the tourism services provided to tourists by measuring the sentiments of tourists in the Twitter hashtags. This paper also evaluates and discusses various research done on tourism data capturing, extraction and analysis. The research approaches used various tourism portals, sites to extract reviews of visitors and the result has been analysed using data analyse tools like SPSS.

This paper proposes a method to analyse the sentiments of tourists using twitter.com social networking site by tweets analysis in active #hashtag and how data is stored in a database as well as enter the keywords that contain the negative and positive words for grouping. The paper also suggests how displaying of result can be done using dashboards and graphs.

The research used live data about the sentiments of visitors on tourist places, which in turn contributes in making the right decisions on tourist places.

V. References

- [1] S. Dutta, T. Geiger, B. Lanvin, "The Global Information Technology Report 2015 ICTs for Inclusive Growth", World Economic Forum, Vol. 1, No. 1, pp.80-85, 2015
- [2] K. A. Manap, "The Role of User Generated Content (UGC) in Social Media for Tourism Sector", WEI International Academic Conference Proceedings, Turkey, pp 52-58, 2013.
- [3] A. Abbasi, T. H. Rashidi, M. Maghrebi, S. T. Waller, "Utilising Location Based Social Media in Travel Survey Methods: bringing Twitter data into the play", Proceedings of the 8th ACM SIGSPATIAL International Workshop on Location-Based Social Networks, USA, pp.1-12, 2015
- [4] R. Turner, "Travel and tourism Economic Impact 2015 – world", World Travel & Tourism Council, United Kingdom, pp.1-24, 2015.
- [5] S. Cerutti, E. Piva, "The Role of Tourists' Feedback in the Enhancement of Religious Tourism Destinations", *International Journal of Religious Tourism and Pilgrimage*, Vol.4, Issue.3, pp. 6-16, 2016.
- [6] K.M. Otto, R L. Cheng, "Big Data Stream Analytics for Near Real-Time Sentiment Analysis", *Journal of Computer and Communications*, Vol.3, Issue.5, pp. 189-195, 2015.

- [7] B. Liu, "Sentiment Analysis and Opinion Mining", Synthesis Lectures on Human Language Technologies, Vol. 5, No. 1, pp.1-167, 2012.
- [8] H. Kwak, C. Lee, H.Park, S.Moon, *What is Twitter, a Social Network or a News Media?*, World Wide Web Conference, USA, 2010.
- [9] A. Bifet, R. Cavada, "Detecting Sentiment Change in Twitter Streaming Data", Proceedings of the Second Workshop on Applications of Pattern Analysis, Spain, pp. 5-11, 2011
- [10] C. Ashley, P. D. Brine, A. Lehr, H. Wilde, "The Role of the tourism sector in expanding economic opportunity", The Role of the tourism sector in expanding economic opportunity, USA, pp. 1-56, 2007.

Authors Profile

Mr. Hamed Saif Albusaidi, High Depo;ma of Computer Science from Sohar University, Oman in 2005. He is currently pursuing Msc-IT. and currently working as I.T Specialist (IT Solutions Developer - web designer). in IT Department, Ministry of Interior, Oman since 2006. He has 10 years of teaching experience and 1 year of Research Experience. self-motivated software developer with a strong technical background who possesses self-discipline and ability to work under the minimum of supervision.

Dr. Prakash has more than 30 years of rich experience in academics, industries, military and training. Dr. Prakash has been involved in various activities related to academics, diversified industries in India and abroad. He was also a former visiting faculty and jury member of various universities and academic institutions. A Chartered Engineer, Corporate Member of IEI and Life Member of QCFI, he is also a distinguished world class trainer on Project Management tools and techniques in India and abroad. As an academician, he also contributed for various academic and industrial publications, journals. His area of research is E-Learning, Multimedia, Computer Graphics, Digital Media, Interactive Media and Project Management. At present, Dr. Prakash Kumar Udipi is a faculty and program manager for Multimedia and Internet Technology specialization at Middle East College, Muscat, Oman.

Dr. Vishal's research/consultancy interest's Mobile ad-hoc networks, Wireless sensor networks, self-healing networks and Network security. He has around 8 years of teaching, research and management experience in Computer science & engineering. Acquired diversified exposure in curricular, co-curricular and extra-curricular activities including institutional development and striving for excellence in technical education and R&D. He is associated with several industries and academic institutions as guest teacher and evaluator.

Dr. Vishal Dattana has received his PhD in Computer Science, and worked at *Global Research Institute of Management and Technology, India*. He is working with Middle East College since year 2013. He has contributed to international conferences as TPC member. He is guiding 2 PhD research scholars as co guide in the area of Networking.