Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.017

IJCSMC, Vol. 7, Issue. 8, August 2018, pg.80 – 91

TWEET ANALYSIS FOR REAL-TIME EVENT DETECTION AND EARTHQUAKE REPORTING SYSTEM

Dr. M. Chidambaram[1], R. Abinaya[2]

Research Guide and Assistant Professor [1], Research Scholar [2]

PG and Research Department of Computer Science, Rajah Serfoji Government College, Thanjavur 613 005, Tamil Nadu – India

ABSTRACT: ***The real time interaction of events such as earthquakes in Twitter and propose an algorithm to monitor tweets and to target events. To detect a target event, devise a classifier of tweets based on features such as the keywords in a tweet, the number of words, and their context. Subsequently, produce a probabilistic spatiotemporal model for the target event that can find the center of the event location. Here regard each Twitter user as a sensor and apply particle filtering, which are widely used for location estimation. The particle filter works better than other comparable methods for estimation the location of target events. As an application, developed an earthquake reporting system for use in japan. Because of the numerous earthquakes and the large number of twitter users throughout the country, propose an event detection algorithm, which helps to detect new events, analyze the spatial and temporal pattern of an event, and to identify importance of events. In this demonstration, the overall system components that crawl, classify, and rank tweets and extract locations from tweets, and present some interesting results of our system.***

1. INTRODUCTION:

Analyze your Tweets and understand your followers. Every word, photo, video, and follower can have an impact. Tweet activity dashboard is where you'll find metrics for every single one of your Tweets. Events usually refer to something abnormal, that is, something that rarely happens in normal situation. [1]

Event detection aims to find such abnormal phenomenon from collected data. Various methods have been proposed to detect events such as disease outbreaks, criticism and so on. When it comes to Twitter, events are topics that suddenly draw public attention, for example, music concerts, football matches and so on. Some topics are not events, although they are popular.

An earthquake's point of initial rupture is called its focus or hypocenter. The epicenter is the point at ground level directly above the hypocenter, a popular micro blogging service, has received much attention recently. This online social network is used by millions of people around the world to remain socially connected to their friends, family members, and co-workers through their computers and mobile phones. Twitter asks one question, "What's happening?" Answers must be fewer than 140 characters. A status update message, called a tweet, is often used as a message to friends and colleagues. A user can follow other users; that user's followers can read her tweets on a regular basis. A user who is being followed by another user need not necessarily reciprocate by following them back, which renders the links of the network as directed. Since its launch on July 2006, Twitter users have increased rapidly. The number of registered Twitter users exceeded 100 million in April 2012. The service is still adding about 300,000 users per day. Currently, 190 million users use Twitter per month, generating 65 million tweets per day. Many researchers have published their studies of Twitter to date, especially during the past year. Most studies can be classified into one of three groups: first, some researchers have sought to analyze the network structure of Twitter. Second, some researchers have specifically examined characteristics of Twitter as a social medium. Third, some researchers and developers have tried to create new applications using Twitter. Twitter is categorized as a micro blogging service. Micro blogging is a form of blogging that enables users to send brief text updates or micro media such as photographs or audio clips. Much attention recently. Researchers have specifically examined characteristics of Twitter as a social medium.

Third, some researchers and developers have tried to create new applications using Twitter. Twitter is categorized as a micro blogging service. Micro blogging is a form of blogging that enables users to send brief text updates or micro media such as photographs or audio clips. Twitter is classified as a microblogging service. Microblogging may be a variety of blogging that allows users to send transient text updates or macromedia like images or audio clips. Microblogging services aside from Twitter embody Tumbler, Friend Feed, characteristic that's common among small blogging services is their period nature. Though web log users usually update their blogs once each many days, Twitter users write tweets many times during a single day.

2. EVENT DETECTION

2.1. Tweets Analysis:

To detect a target event from Twitter, we search from Twitter and find useful tweets. Our method of acquiring useful tweets for target event detection. Tweets might include mention of the target event. For example, users might make tweets such as “Earthquake!” or “Now it is shaking.” Consequently, earthquake or shaking might be keywords (which we call query words). However, users might also make tweets such as “I am attending an Earthquake Conference.” Or “Someone is shaking hands with my boss.” Moreover, even if a tweet is referring to the target event, it might not be appropriate as an event report. For instance, a user makes tweets such as “The earthquake yesterday was scary.” or “Three earthquakes in four days. Japan scares me.” These tweets are truly descriptions of the target event, but they are not real-time reports of the events. Therefore, it is necessary to clarify that a tweet is truly referring to an actual contemporaneous earthquake occurrence, which is denoted as a positive class. To classify a tweet as a positive class or a negative class, we use a support vector machine, which is a widely used machine-learning algorithm. By preparing positive and negative examples as a training set, we can produce a model to classify tweets automatically into positive and negative categories. [1]

2.2.Tweet as a Sensory Value:

The tweet and classify it into a positive class if a user makes a tweet about a target event. In other words, the user functions as a sensor of the event. If she makes a tweet about an earthquake occurrence, then it can be considered that she, as an "earthquake sensor," returns a positive value. A tweet can therefore be regarded as a sensor reading. This crucial assumption enables application of various methods related to sensory information. The sensors are vastly numerous: there are more than 100 million 'Twitter sensors' worldwide producing tweet information around the clock. A sensor might be inoperable or operating incorrectly sometimes (which means a user is not online, sleeping, or is busy doing something else). For that reason, this social sensor is noisier than ordinary physical sensors such as location sensors, thermal sensors, and motion sensors.

3. MODULES [5]

- Tweet collection Module
- Crawling tweets from Twitter Module
- Twitter Search API Module
- Filtering tweets using machine learning Module
- Semantic Analysis on Tweets Module
- Earthquake reporting System Module

3.1. Tweet Collection Module:

In this module, develop our system by posting tweets by the users. It is necessary to collect tweets referring to an earthquake from Twitter. This process includes two steps: crawling tweets from Twitter and filtering out tweets that do not refer to the earthquake. For crawling and filtering tweets, we recommend using script programming languages.

3.2. Crawling Tweets from Twitter Module

To collect tweets or some user information from Twitter, one must use the Twitter Application Programmers Interface (API). Twitter API is a group of commands that are necessary to extract data from Twitter. Twitter has APIs of three kinds: Search API, REST API, and Streaming API. In this section, we introduce Search API and Streaming API, which are necessary to crawl tweets from Twitter. We explain REST API later because REST API is

necessary to extract location information from Twitter information. Additionally, it is known that Twitter API specifications are subject to change. When using Twitter API, it is necessary to know the latest details and requirements. They are obtainable from Twitter API documentation

3.3. Twitter Search API Module:

The Twitter Search API extracts tweets from Twitter, including search keywords or those fitting other retrieval conditions, in chronological order. It is possible to use language, date, location and other conditions as retrieval conditions. Some points must be considered when using Twitter Search API: It is possible to collect tweets posted only during the prior five days. It is not possible to search tweets posted six days ago. It is only possible to collect the latest 1500 tweets at one time. (Technically speaking, it is possible to access one page with a request and track pages back to the 15th page. One page includes 100 tweets at most. Therefore it is possible to acquire the latest 1500 tweets at one time.) One is limited to API requests.

3.4. Filtering Tweets using Machine Learning Module:

Here collected data from tweets including keywords related to earthquakes, such as earthquake, shake. Those tweets include not only tweets that users posted immediately after they felt earthquakes, but also tweets that users posted shortly after they heard earthquake news, or perhaps they misinterpreted some sense of shaking from a large truck passing nearby.

3.5. Semantic Analysis on Tweets Module:

Semantic Analysis on Tweet Search tweets including keywords related to a target event Example: In the case of earthquakes “shaking”, “earthquake” Classify tweets into a positive class or a negative class Example: “Earthquake right now!!” Positive “Someone is shaking hands with my boss” negative Create a classifier

3.6. Earthquake Reporting System Module:

In this module, the users will be altered if the earthquake occurs based on their location and the tweets. Effectiveness of alerts of this system Alert E-mails urges users to prepare for the Earthquake if they are received by a user shortly before the earthquake actually arrives.

4. EARTHQUAKES:

2017 Tripura earthquake:

A magnitude 5.7 earthquake struck India 20 km (12 mi) east north-east of Amass in the state of Tripura on January 3, 2017 with a maximum observed intensity of 6-7 EMS. It struck at 2:39 pm local time (09:09 UTC), and was centered in an isolated area. The estimated depth was 32.0 km. One person died and five others were injured in India. At least 50 houses were damaged due to landslides that occurred in Dahlia district, while roads were blocked after trees were uprooted. According to the Tripura State Disaster Management Authority, at least 6,727 buildings were damaged in Tripura in the districts of Dahlia and Unicoi. Shaking was felt in many parts of north-eastern India including as far as Kolkata. The tremor was also felt in neighboring Bangladesh, where two people died and three others were injured. The earthquake caused liquefaction on the banks of the Manu River in Tripura and along the Dahlia River in adjacent parts of Bangladesh, in particular in the Amalgam area.

2016 Impala earthquake:

The 2016 Impala Earthquake struck northeast India in the state of Manipur on January 4 with a moment magnitude of 6.7 and a maximum Marcella intensity of VII (Very strong). At least eleven people were killed, 200 others were injured and numerous buildings were damaged. The quake was also strongly felt in Bangladesh. It was also extensively felt in eastern and north-eastern India. The Earthquake, which hit at 4:35 a.m. on 4 January local time (23:05 UTC, 3 January), was centered in an isolated area. Impala has a population of more than 250,000. It was one of the most damaging Earthquakes in Manipur since 1880 and 1939.

2015 Hindu Kush earthquake:

The October 2015 Hindu Kush earthquake was a magnitude 7.5 Earthquake that struck South Asia on 26 October 2015, at 13:39 AFT (14:09 PKT; 14:39 IST; 09:09 UTC) with the epicenter 45 km north of `Alaqahdari-ye Karan we Mundane, Afghanistan, at a depth of 212.5 km. By 5 November, it was estimated that at least 399 people were killed, mostly in Pakistan. Tremors were felt in Pakistan, Uzbekistan, Turkmenistan, Tajikistan and Kyrgyzstan. According to the United States Geological Survey (USGS), Pakistan is located in one of the most Earthquake active zone in the world. The earthquake was also felt in New Delhi, in both Pakistan occupied Kashmir and Indian state of Jammu and Kashmir and as far as Luck now and in the

prefectures of Kasha, Aksum, Hogan, and Kisi's in Xinjiang, China while damage was also reported in Afghan capital Kabul. The earthquake was also felt in the Nepalese capital of Kathmandu.

2011 Sikkim earthquake:

The 2011 Sikkim earthquake (also known as the 2011 Himalayan earthquake) occurred with a moment magnitude of 6.9 and was centered within the Kanchenjunga Conservation Area, near the border of Nepal and the Indian state of Sikkim, at 18:10 IST on Sunday, 18 September. The earthquake was felt across northeastern India, Nepal, Bhutan, Bangladesh and southern Tibet. At least 111 people were killed in the earthquake. Most of the deaths occurred in Sikkim, with reports of fatalities in and near Sing tam in the East Sikkim district. Several buildings collapsed in Bangkok. Eleven are reported dead in Nepal, including three killed when a wall collapsed in the British Embassy in Kathmandu. Elsewhere, structural damage occurred in Bangladesh, Bhutan, and across Tibet; another seven fatalities were confirmed in the latter region. The quake came just a few days after an earthquake of 4.2 magnitude hit Haryana's Sonata district, sending tremors in New Delhi. The Earthquake was the fourth significant earthquake in India of September 2011. Exactly a year after the original earthquake at 5:55 pm on 18 September 2012, another earthquake of magnitude struck Sikkim, sparking panic among the people observing the anniversary of the original quake.

2009 Andaman Islands earthquake:

The 2009 Andaman Islands earthquake occurred on August 11 at 19:55 UTC in the Andaman Islands of India. The earthquake magnitude was recorded as 7.5 Mw, and was the strongest in the region since the 2004 Indian Ocean earthquake and tsunami. The epicenter was 260 km north of Port Blair, and tremors were felt in south-east India, Bangladesh, Myanmar, and Thailand. The Pacific Tsunami Warning Center issued a tsunami watch to India, Myanmar, Bangladesh, Indonesia and Thailand, but it was later lifted. No casualties or injuries were reported, although there were complaints about minor damage to buildings. At about the same time, the 2009 Shizuoka earthquake affected south Honshu in Japan.

2005 Kashmir earthquake:

The 2005 Kashmir earthquake occurred at 08:50:39 Pakistan Standard Time on 8 October in Pakistan administered areas of Kashmir. It was centered near the city of Muzaffarabad, and

also affected Pakistan's Khyber Pakhtunkhwa province and the Indian State of Jammu and Kashmir. It registered a moment magnitude of 7.6 and had a maximum Marcella intensity of VIII (Severe). The earthquake also affected countries in the surrounding region where tremors were felt in Afghanistan, Tajikistan and Chinese Xinjiang. The severity of the damage caused by the earthquake is attributed to severe up thrust. It is considered the deadliest earthquake to hit South Asia since the 1935 Quetta earthquake.

4.1. Prediction methods:

Earthquake prediction is an immature science. It has not yet led to a successful prediction of an Earthquake from first physical principles. Research into methods of prediction therefore focus on empirical analysis, with two general approaches: either identifying distinctive precursors to earthquakes, or identifying some kind of geophysical trend or pattern in seismicity that might precede a large earthquake. Precursor methods are pursued largely because of their potential utility for short-term Earthquake prediction or forecasting, while 'trend' methods are generally thought to be useful for forecasting, long term prediction (10 to 100 years' time scale) or intermediate term prediction (1 to 10years time scale).

4.1.1 Precursors

An Earthquake precursor is an anomalous phenomenon that might give effective warning of an impending earthquake. Reports of these – though generally recognized as such only after the event – number in the thousands, some dating back to antiquity. There have been around 400 reports of possible precursors in scientific literature, of roughly twenty different types, running the gamut from agronomy to zoology. None have been found to be reliable for the purposes of earthquake prediction. In the early 1990, the IASPEI solicited nominations for a Preliminary List of Significant Precursors. Forty nominations were made, of which five were selected as possible significant precursors, with two of those based on a single observation each. After a critical review of the scientific literature the International Commission on Earthquake Forecasting for Civil Protection (ICEF) concluded in 2011 there was "considerable room for methodological improvements in this type of research". In particular, many cases of reported precursors are contradictory, lack a measure of amplitude, or are generally unsuitable for a rigorous statistical evaluation. Published results are biased towards positive results, and so the rate of false negatives (earthquake but no precursory signal) is unclear.

4.1.2. Animal behavior

For centuries there have been anecdotal accounts of anomalous animal behavior preceding and associated with earthquakes. In cases where animals display unusual behavior some tens of seconds prior to a quake, it has been suggested they are responding to the P-wave. These travel through the ground about twice as fast as the S-waves that cause most severe shaking. They predict not the earthquake itself — that has already happened — but only the imminent arrival of the more destructive S-waves. It has also been suggested that unusual behavior hours or even days beforehand could be triggered by foreshock activity at magnitudes that most people do not notice. Another confounding factor of accounts of unusual phenomena is skewing due to "flashbulb memories": otherwise unremarkable details become more memorable and more significant when associated with an emotionally powerful event such as an earthquake. A study that attempted to control for these kinds of factors found an increase in unusual animal behavior (possibly triggered by foreshocks) in one case, but not in four other cases of seemingly similar earthquakes.

5. RESULT AND DISCUSSIONS

The following table shows the earthquake comparative performance of the above discussed magnitude.

YEAR	MAGNITUDE	DEATHTOLL	LOCATION	DATE
2018	7.9	0	US	JAN 23
2017	8.2	98	MEXICO	SEP 3
2016	7.8	676	ECUADOR	APR 16
2015	8.3	14	CHILE	SEP 16
2014	8.2	6	CHILE	APR 1
2013	8.3	0	RUSSIA	MAY 24
2012	8.6	10	INDONESIA	APR 11
2011	9.1	20896	JAPAN	MAR 11
2010	8.8	525	CHILE	FEB 27
2009	8.1	192	SAMOA	SEP 29

2008	7.9	87587	CHINA	MAY 12
2007	8.5	23	INDONESIA	SEP 12
2006	8.3	0	RUSSIA	NOV 15
2005	8.6	1300	INDONESIA	MAR 28
2004	9.2	227898	INDONESIA	DEC 26
2003	8.3	0	JAPAN	SEP 25
2002	7.9	0	US	NOV 3

Table-1: Magnitude value

5.1. LARGEST EARTHQUAKE BY MAGNITUDE:

Listed below are all the 36 known earthquakes with an estimated magnitude of 8.5 or higher since the year 1500. Limited to a timeframe with enough data, this gives a rough estimate of its frequency per century. (The timeframe does not include outlying events like the earlier 1361 Shyheim earthquake and 869 Shanika earthquake, both estimated to have magnitude ≥ 8.5.)

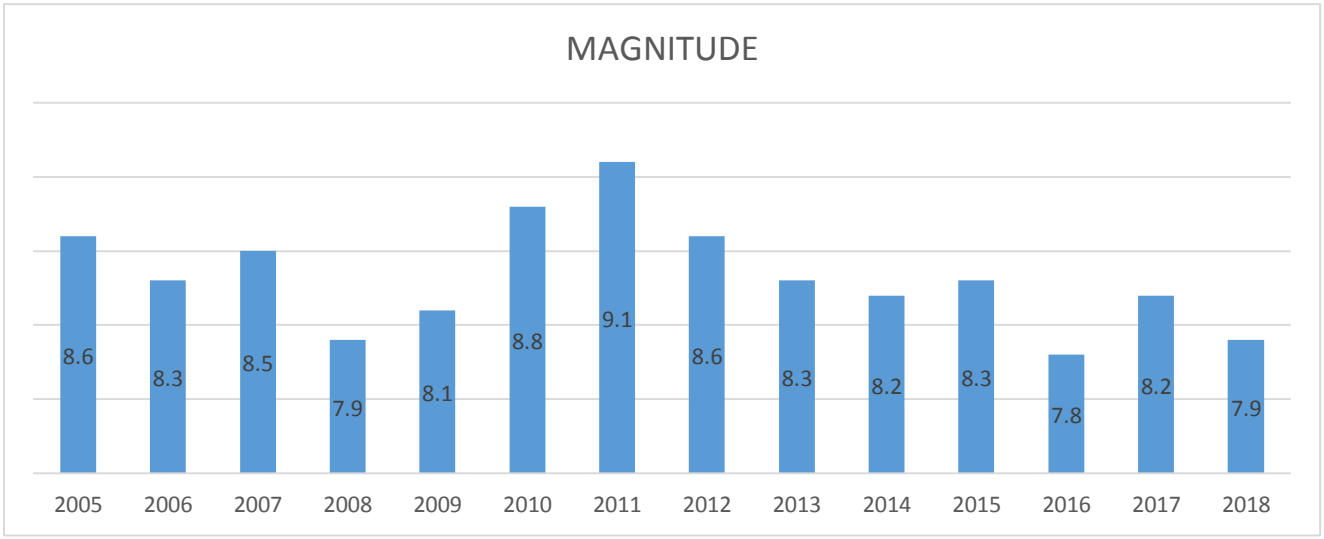

Figure-: Largest Magnitude

6. CONCLUSION:

As described in this paper, investigated the real-time nature of Twitter, devoting particular attention to event detection. Semantic analyses were applied to tweets to classify them into a positive and a negative class. Microblogging has real-time characteristics that distinguish it from other social media such as blogs and collaborative bookmarks. As described in this paper, here presented an example that leverages the real-time nature of Twitter to make it useful in solving an important social problem: natural disasters. It is hoped that this paper will provide some insight into the future integration of semantic analysis with microblogging data. Real time event detection system aims at detecting some events in real time and take the required actions based on the detection. Our project "Real time event detection of earthquake using twitter " aimed at building a similar system ,concluded at providing an estimate to the tweet that is input to the system. The proposed model can be used in real time with some modifications to help people act in case of earthquake. The project topic demands further research in future.

REFERENCES:

1. X. Liu, S. Zhang, F. wee and M. Zhou. "Recognizing Named Entities in tweets" Jun 2011
2. A. Ritter, S. Clark Maksim and O. E-zine, " Named Entity Recognition in Tweets: An Experimental study" 2011
3. A. Ritter, S. Clark Maksim, and O. E-zine, "Named Entity Recognition in Tweets: An Experimental Study," Proc. Conf. Empirical Methods in Natural Language Processing, 2011.
4. H. Kwan, C. Lee, H. Park, and S. Moon, "What is Twitter, A Social Network or A News Media?" Proc. 19th Int'l Conf. World Wide Web (WWW '10), 2010.
5. G.L. Dinah Boyd and S. Gilder, "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter," Proc. 43rd Hawaii Int'l Conf. System Sciences (HICSS-43), 2010.
6. A. Timespan, T.O. Springer, P.G. Sander, and I.M. Whelped, "Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment," Proc. Fourth Int'l AAAI Conf. Weblogs and Social Media (ICWSM), 2010.

7. P. Gallegan, "Twitter as a Learning Tool. Really," ASTD Learning Circuits, p. 13, 2009
8. Twitter search for Haiti Survivors channel 4. 15 Jan,2010
9. T. Sakami, M. Okazaki, and Y. Matsuo, "Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors," Proc. 19th Int'l Conf. World Wide Web (WWW '10), pp. 851-860, 2010
10. A survey of techniques for event detection in twitter.

