American Journal of Engineering Research (AJER) 2016
American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-5, Issue-2, pp-199-203
www.ajer.org
Research Paper Open Access

A Survey Paper on Real Time Detection & Reporting By Social Networks

Thokal Nisha E., Golekar Sulbha S., Vavhal Priyanka B., Ashish Kumar
Department of CSE, G.H.Raisoni College of Engg. & Management, Ahmednagar, Savitribai Phule Pune University, India,414005.

***ABSTRACT*:** *In today's era social network like Twitter, Facebook, LinkedIn playing very vital role in human behavior & their day to day life. These social networking sites can be used as social sensors such that we can predict or report Real time natural disastrous events likeStorms, fire, traffic jams, riots, heavy rain-falls, earthquakes. And also social events Parties, baseball games, presidential campaign .We have used twitter here for such purpose to have a tendency to investigate the period of time interaction of events like earthquakes in Twitter associate degreed propose a rule to observe tweets and to sight a target event. To sight a target event, we have a tendency to devise a classifier of tweets supported options like the keywords in a very tweet, the amount of words, and their context. This model is provided which might notice the centre of the event location & it will do real time reporting to all active users. The twitter users also act as a sensors and report the event information by making tweet or post. Once our module gets such tweets then it will start analysis and ready to predict real time event & location. Our reporting system also considers amount of tweets generated from same region so that system can check the authenticity of event. We choose earthquake related datasets for training at initial as we found it is very essential and continuously occurring event all over the world. We also can change our module with respect to different data sets of different events.*

***Keywords*:** *Twitter, event detection, social sensing element, location estimation, Earthquake, Hadoop- Data processing, Machine Learning, Probabilistic Model, Earthquake monitoring, Instruments and techniques, Seismological data, Social Media, General or miscellaneous.*

I. INTRODUCTION

Twitter is an online social networking service that enables users to send and read short 180 character messages called "tweets". Registered users can read and writes tweets, but those who are not registered can only read . Users access Twitter through the website interface, SMS or mobile device app[19]. The service rapidly gained worldwide popularity, with more than 100 million users posting 340 million tweets a day in 2012.The service also handled 1.6 billion search quries per day. In 2013, Twitter was one of the ten most visited websites and has been described as "the SMS of the Internet".As of May 2015, Twitter has more than 500 million users, out of which more than 302 millions are active. Twitter is classified as a microblogging service. Microblogging may be a variety of blogging that allows users to send transient text updates or micromedia like images or audio clips[15]. An important 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.In such a way, varied update leads to varied reports associated with events. They embrace social events like parties, baseball games, and presidential campaigns. They additionally embrace black events like storms, fires, traffic jams, riots, significant precipitation, and earthquakes. Actually, Twitter is employed for varied period of time notifications like that necessary for facilitate throughout a large-scale hearth emergency or live traffic updates. The analysis question of our study is, "can we tend to discover such event prevalence in period of time by watching tweets?" This paper presents associate investigation of the period of time nature of Twitter that's designed to determine whether or not we are able to extract valid data from it. we tend to propose an occurrence notification system that monitors tweets and delivers notification promptly victimisation information from the investigation. As an application, we develop an earthquake reporting system using Japanese tweets. Japan has numerous earth- quakes. Twitter users are similarly numerous and geogra- phically dispersed throughout the country. Therefore, it is sometimes possible to detect an earthquake by monitoring tweets. Our system detects an earthquake occurrence and sends an e-mail, possibly before an earthquake actually arrives at a certain location: An earthquake propagates at about 3-7 km/s. For that reason, a person who

is 100 km distant from an earthquake is able to communicate and act for about 20 s before the arrival of an earthquake wave. Moreover, strong earthquakes often cause tsunami, which engender more catastrophic disasters than the earthquakes themselves in distant and near places in relation to the earthquake epicenter, as did the Haiti earthquake in 2010 and the Great Eastern Japan earthquake in 2011. Therefore, prompt notification of earthquake occurrences is extremely important to decrease damage by tsunami. In many cases, it could provide notification of tens of minutes or even hours before a tsunami strikes a coastal area[1].

II. RELATED WORK

2.1 EVENT DETECTION: A survey[1] by Takeshi Sakaki, Makoto Okazaki, Yutaka Matsuo,2013
As described in this paper, we target event detection.We target events such as earthquakes, typhoons, and traffic jams, which are readily apparent upon examination of tweets. These events have several properties.

1.many users experience the event. 2.for that reason, people are induced to tweet about it. 3.so that real-time location estimation is possible. Such events include social events such as large parties, sports events, exhibitions, accidents, and political cam- paigns. They also include natural events such as storms, rain are heavy, tornadoes, typhoons/hurricanes/cyclones, and earthquakes. We designate an event we would like to detect using Twitter as a target event.

In this section, we explain how to detect a target event using Twitter. First, we crawl tweets including keywords related to a target event. From them, we extract tweets that certainly refer to a target event using devices that have been trained with machine learning. Second, we detect a target event and estimate the location from those tweets by treating Twitter users as "social sensors."

2.2 Semantic Analysis of Tweets:

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 is portrayed in. 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 occur- rence, 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 [14], 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.
We prepare three groups of features for each tweet
statistical features: the number of words in a tweet message, and the position of the query word within a tweet.
keyword features: the words in a tweet.
word context features: the words before and after the query word.
We can give an illustrative example of these features using the following sentence. "I am in Japan, earthquake right now!"
keyword: earthquake
Using the obtained model, we can classify whether a new tweet corresponds to a positive class or a negative class.

2.3 Semantic and Sentiment Analysis:[2] Prof Shivani Desai,Priyank Bhatt,Vraj Solanki,Sep 2014.
This paper aims at surveying a number of such algorithms, methods and techniques to classify any sentence as positive, negative or neutral and also discuss the issues related to each method faced during implementation and execution. The essential issues in sentiment analysis are to identify how sentiments are expressed in texts and whether the expressions indicate positive (favorable) or negative (unfavorable) opinions toward the subject and how efficiently and correctly sentences are classified.

2.4 SYSTEM OVERVIEW

In our model we are focusing on submodules like Tweeter dataset crawling for training purposes. Then Implementing machine learning module and data analysis module on Hadoop framework is very important part. Interactive GUI which gave us real time analysis reports and charts is last part of project where active users will get notifications for such events. We tend to are getting to propose an occurrence notification system. An occurrence watching system monitors tweets and delivers notification promptly mistreatment investigation results. We tend to propose a system that's supported investigation of tweets i.e. real time investigation. during this analysis, in brief, we tend to take 3 steps:

1) We analyze no of tweets associated with target events;

2) We got to style such a probabilistic module to research and extract events from those tweets and predict locations of events with category verifying as positive and negative class.

3) Finally developed coverage system that extracts earthquakes from Twitter and sends a message to registered users.

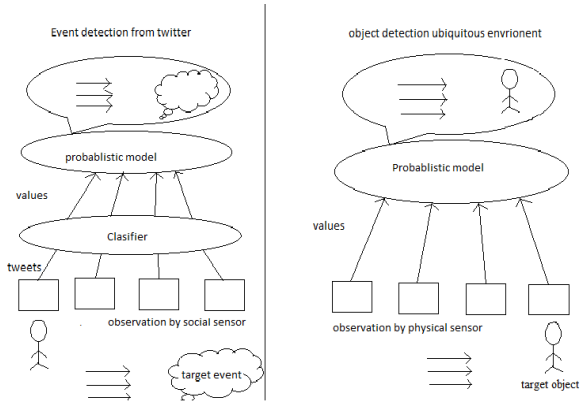

Fig 1:System overview

2.5 IMPLEMENTATION:[3]S.Anand & K.Narayana

This paper presents an investigation of the real-time nature of Twitter that is designed to ascertain whether we can extract valid information from it. We propose an event notification system that monitors tweets and delivers notification promptly using knowledge from the investigation. In this research, we take three steps: first, we crawl numerous tweets related to target events; second, we propose probabilistic models to extract events from those tweets and estimate locations of events; finally, we developed an earthquake reporting system that extracts earthquakes from Twitter and sends a message to registered users.

2.5 METHODS

For event detection and placement estimation, we tend to use probabilistic models. during this section, we tend to 1st describe event detection from time-series information. Then we tend to describe the situation estimation of a target event.

1)Temporal Model

Each tweet has its own post time. once a target event happens, however do the sensors discover the event, we tend to describe the temporal model of event detection. First, we tend to examine the particular information. The several quantities of tweets for a target event: Associate in Nursing earthquake. it's apparent that spikes occur within the variety of tweets. every corresponds to an incident} occurrence. Specifically concerning Associate in Nursing earthquake, quite ten earthquakes occurred throughout the amount.

2) Spatial Model

Each tweet is related to a location. we tend to describe a technique which will estimate the situation of an occasion from device readings. to resolve the matter, many ways of Bayesian filters square measure planned like Kalman filters, multi-hypothesis following, grid-based and topological approaches, and particle filters. For this study, we tend to use particle filters, each of that square measure wide employed in location estimation.

a) Particle Filters

A Particle filter could be a probabilistic approximation algorithmic rule implementing a Bayes filter, and a member of the family of successive Monte Carlo strategies.

b) Consideration of sensing element Geographic Distribution

We should take into account the sensing element geographic distribution to treat readings of social sensors additional. exactly. In location estimation by physical sensors, those sensors area unit situated equally in several cases. we will treat sensing element readings equally in such things. Actually, social sensors aren't placed equally in several cases as a result of social media user's area unit targeted in urban areas. In Japan, most users board capital of Japan. Therefore, we should always incorporate the geographic distribution of social sensors into abstraction models.

c) Techniques to hurry up the method

As represented during this paper, we wish to estimate location of events quickly as shortly as potential as a result of one objective of this analysis is to develop a period earthquake detection system. Therefore, we tend to should decrease the time quality of strategies used for location estimation.

3) Information Diffusion associated with a period Event

Some info associated with an occasion diffuses through Twitter. for instance, if a user detects associate earthquake and makes a tweet regarding the earthquake, then a fan of that user would possibly create tweets that. This characteristic is very important as a result of, in our model, sensors won't be reciprocally freelance, which might have associate unsought result in terms of event detection.For event detection and placement estimation, we tend to use probabilistic models. From time-series information, we 1st describe event detection. Then we tend to describe the placement estimation of a target event. Each tweet has its own post time. once a target event happens, however do the sensors observe the event? we tend to describe the temporal model of event detection. First, we tend to examine the particular information. every corresponds to prevalence occurrence. Specifically relating to associate earthquake, over ten earthquakes occurred throughout the amount.

2.6 PROPOSED WORK:

We are going to design the system called 'CrisisAlertCall' is kind of alarming or reporting system based on Hadoop Framework to process huge amount of tweeter data related with 'Earthquake' like calamities. Actually it is one of much needed project started by seeing hazards to people in Nepal. Recently in July-August 2015 Nepal faced very big natural calamity due to Earthquake. Many people lost their lives. By seeing such huge hazard we are proposing such system which will make reports or alarms to social network users on their accounts or public announcing on web by analyzing huge amount of Tweets, Posts related with. 'CrisisAlertCall' is prepared by not only focusing on Earthquake, but also events like Storm, Heavy Rainfall, Flood etc. can be monitor by us. One of best thing in this project is we are using Hadoop Framework which was not present in existing system. Means we can process very huge data within very short time period.

Our Systems' flow will starts from Tweets crawling. First we collect as many tweets as possible from Tweeter Database or else we crawl the web for Tweets. Then we will process them as Positive or Negative Tweets using Navie Bays Algorithm. So we have to make them sorted as per feeling or sentiment in tweet. In such scenarios we will use sentiment analysis concepts do make type sorting of tweets. Also from such tweets we have detect event and location too. So it is one of big task while developing. As tweeter data gives us locations as time of tweet we can Processed data of Messenger is used for preparing the algorithm, so that next time if such data came for prediction it will processed directly. For processing we will use Navie bays algorithm over Hadoop Map Reduce framework. When real time Tweeter users tweet on web our system detects event & location & does reporting & alarming automatically. 'CrisisAlertCall' is best prototype for Harmful event detection & location tracking for any kind. What just change is we need to change database for training the system & evaluation.

III. CONCLUSION

we have a tendency to investigated the period nature of Twitter, devoting specific attention to event detection. linguistics analyses were applied to tweets to category verify them into a positive and a negative class. we have a tendency to regard every Twitter user as a device, and set the matter as detection of a happening supported sensory observations. Location estimation strategies like particle filtering area unit used to estimate the locations of events. As associate degree application, we have a tendency to developed associate degree earthquake coverage system, that could be a novel approach to advice folks promptly of associate degree earthquake event.

REFERENCES

[1] T. Sakaki, M. Okazaki, and Y. Matsuo, "*Tweet Analysis for Real-Time Event Detection and Earthquake Reporting System Development,*" Proc. Int'l Conf. World Wide Web (WWW '10), pp. 9526, Vol 25 April 2013.

[2] Prof Shivani Desai, Priyank Bhatt(11BIT044), Vraj Solanki(11BIT059) Semantic and Sentiment Analysis,sep2014.

[3] S.Anand, K.Narayan," *Earthquake Reporting System Development by Tweet Analysis*" IJEERT Volume 2, Issue 4, July 2014, PP 96-106 ISSN 2349-4395 (Print) & ISSN 2349-4409

[4] B. Huberman, D. Romero, and F. Wu, "*Social Networks that Matter: Twitter Under the Microscope,"* ArXiv E-Prints, http://arxiv.org/abs/0812.1045, Dec. 2008.

[5] H. Kwak, 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), pp. 591-600, 2010.

[6] G.L. Danah Boyd and S. Golder, "*Tweet, Tweet, and Retweet: Conversational Aspects of Retweeting on Twitter,"* Proc. 43rd Hawaii Int'l Conf. System Sciences (HICSS-43), 2010.

[7] Tumasjan, T.O. Sprenger, P.G. Sandner, and I.M. Welpe, *"Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment,"* Proc. Fourth Int'l AAAI Conf. Weblogs and Social Media (ICWSM), 2010.

[8] P. Galagan, *"Twitter as a Learning Tool. Really,"* ASTD Learning Circuits, p. 13, 2009.

[9] K. Borau, C. Ullrich, J. Feng, and R. Shen, *"Microblogging for Language Learning: Using Twitter to Train Communicative and Cultural Competence,"* Proc. Eighth Int'l Conf. Advances in Web Based Learning (ICWL '09), pp. 78-87, 2009.
[10] J. Hightower and G. Borriello, *"Location Systems for Ubiquitous Computing,"* Computer, vol. 34, no. 8, pp. 57-66, 2001.
[11] M. Weiser, *"The Computer for the Twenty-First Century,"* Scientific Am., vol. 265, no. 3, pp. 94-104, 1991.
[12] V. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "*Bayesian Filtering for Location Estimation*," IEEE Pervasive Computing, vol. 2, no. 3, pp. 24-33, July-Sept. 2003.
[13] T. Sakaki, 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.
[14] M. Sarah, C. Abdur, H. Gregor, L. Ben, and M. Roger, "*Twitter and the Micro-Messaging Revolution*," technical report, O'Reilly Radar, 2008.
[15] *Microblogging – Wikipedia, the free encyclopedia.*
[16] Java, X. Song, T. Finin, and B. Tseng, "*Why We Twitter: Understanding Microblogging Usage and Communities,"* Proc. Ninth WebKDD and First SNA-KDD Workshop Web Mining and Social Network Analysis (WebKDD/SNA-KDD '07), pp. 56-65, 2007.
[17] *Twitter search for Haiti survivors Channel 4 15 Jan, 2010.*
[18] Krūms, Jānis (January 15, 2009). *"There's a plane in the Hudson. I'm on the ferry going to pick up the people.*
[19] Paul S. Earle*, Daniel C. Bowden, Michelle Guy(ANNALS OF GEOPHYSICS, 54, 6, 2011; doi: 10.4401/ag-5364)"Twitter earthquake detection: earthquake monitoring in a social world".