TWEET ANALYSIS FOR REAL-TIME NATURAL CALAMITY

Mr. Vineet Jadhav, Mr. Ganesh Shinde, Mr. Ganesh Bhakare

[1] *Mr. Vineet Jadhav, Computer Engineering, Pravara Rural Engineering College Loni, Maharashtra, India*
[2] *Mr. Ganesh Shinde, Computer Engineering, Pravara Rural Engineering College Loni, Maharashtra, India*
[3] *Mr. Ganesh Bhakare, Computer Engineering, Pravara Rural Engineering College Loni, Maharashtra, India*

ABSTRACT

Twitter has much popular recently. It has attracted number of users to share the information and communication to each other. It produces the large amount of data every day. An important characteristic of Twitter is its real-time nature. The system investigates the real-time interaction of events such as natural calamity using Twitter and proposes an algorithm to control tweets and to detect a target event. To detect a target event, system divides a classifier of tweets depends on features such as the keywords in a tweet, the number of words, and their context. Subsequently, system to form a probabilistic spatiotemporal model for the target event that can find the center of the event location. System regard every Twitter user as a sensor and apply particle filtering, which are widely used for location estimation. The particle filter works better than other various methods for estimating the locations of target events. As an application, system develops the natural calamity reporting system for use in no of country. Because of the number of natural calamity and the so many number of Twitter users throughout the country, system can detect the natural calamity with high probability. Merely by monitoring tweets, Our system detects natural the calamity promptly and notification is delivered much faster.

Keyword : - *Text Mining , Tweets, Event Detection, and Location Estimation etc….*

1. Introduction

Today Social Networking Sites (SNS) have become a part of our day to day life. We share a lot of data on these sites. They helped us to make the world smaller and integrated with each other. There are many SNS available today and many more are piling each day. Thus users use many SNS each day and communicate and share data with friends and family. This communication medium gave rise to complex structure whether a user really like the SNS which he uses more or he needs another SNS other than he uses more. Thus one of the most famous SNS is TWITTER which is used to share data and post our thoughts and latest buzz upon the internet. The users using TWITTER have increased dramatically in the recent years. So the analysis of this SNS may help in answering and predicting many answers. This online social network (twitter, Facebook, etc.) is used by millions of people around the world to remain socially connected to their friends, family members, and work met through their computers and mobile phones. Answer is less than 140 characters when Twitter asks the question, "What's happening?". A status update message, called a tweet, is often used as a tweet to friends and colleagues. A one user can follow other users; that user's followers can read her tweets on a daily 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 fast. The number of registered Twitter users added 100 million in April 2010. The service(twitter) is still adding about 300,000 users per day.1 now a days, 190 million users use Twitter per month, generating 65 million tweets per day.

2. Programming code

JAVA programming codes for building application to analyze the real time event detection like earthquake, flood , etc.

3. Tools for development

To develop the application, windows xp or higher operating systemis required. Java development kite (Jdk 1.6) or higher. TWITTER REST API Is use to monitor the tweets.

4. Methodology

4.1 Support vector machines-1

SVM are supervised learning models which analyze data and recognize patterns associated with learning algorithms that, used for classification and regression analysis. The basic SVM takes a set of input data and predicts, for input given to system, which of two possible classes forms the output, forming it a non-probabilistic binary linear classifier.

4.2 The Twitter Search API-2

The Twitter Search API is part of Twitter's v1.1 REST API. Recent or popular Tweets allows as queries against the indices of and behaves similarly to, but not exactly like the Twitter mobile or web clients Search feature available, such as Twitter.com search.

5. LITERATURE SURVEY

We have been gathering information about how to integrate the twitter with our application and how to monitor tweets using twitter rest API. For analyzing and predicting tweets we are thinking of using machine learning algorithms as they are strong enough to predict positive and negative results after they are trained and the proposed system will become user independent. We are thinking of using google maps to show the results properly. So for achieving those we are presenting the material needed for our application to perform properly.

6. How it works

6.1 Detecting Events-1

As Twitter is a real-time entity, the different events can be detected by analysing the tweets. Spatial and temporal information, about tweet is present from that we can detect time and location. E.g., via monitoring an incoming tweet "Shooting outside inorbit cafe." at 2:38pm on July 24, we were able to detect the crime immediately, and extract the location and the time of the event as well.

6.2 Analysing Events-2

Twitter gives different reports for the same event (e.g., "Japan earth quake"), by collecting tweets for that event and analyze aspects for that information. First, we observe patterns the spatial pattern and temporal pattern of events. E.g., via analyzing the distribution of tweets related event, we can observe the major regions for that event. Second, we can identify the time-period of important events in that region. E.g., via analyzing the number of tweets related to Japan earthquake, identify that event as an influential event worldwide. Recently, several systems (e.g., [1]) have been proposed to observe tweets and detect events, but most of them are missing the analysis component.

In the literature, systems are proposed to analyze events from blogs, but they may fail in processing tweets, which are short and noisy, and do not give the correct information (e.g., user's network) in Twitter.

7. Architecture

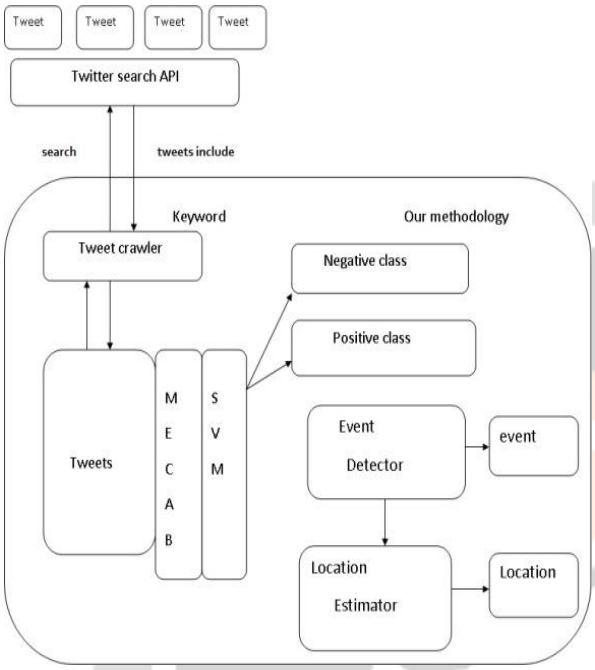

Fig. : Architecture of Tweet analysis for real–time natural calamity

8. Security issues

To make the system secure the user need to know the password of the application. Log file should be maintained to keep track of the use of the application.

9. Advantages

1: The propose system is user friendly and have search tweet mechanism.

2: The system is used for prediction of destination location.

3: The system helps in finding exact location of natural disasters by classifying the tweets.

4: Results should be analyzed properly by using SVM classifier.

5: The system will analyze the results and show on Google map.

10. CONCLUSIONS

As described in this paper, we investigated the real-time nature of Twitter, devoting particular attention to event detection. Semantic analyses were performing, on tweets to divide them into a positive and a negative class. We regard each Twitter user as a sensor, and set the problem as detection of an event based on sensory observations. Location estimation methods such as classify and filtering are used to find the locations of events. As an application, we developed an natural event system, which is a novel approach to inform people promptly of an earth-quake event. Micro blogging has real-time characteristics that is look different it from other social media such as blogs and

collaborative bookmarks. As described in this paper, we give an example that leverages the real-time nature of Twitter to make it useful in solving an very important social problem: natural disasters. It is hoped that this paper will provide some insight into the time ahead integration of semantic analysis with micro blogging data.

11. REFERENCES

[1] M. Sarah, C. Abdur, H. Gregor, L. Ben, and M. Roger, "Twitter and the Micro-Messaging Revolution," technical report, O'Reilly Radar, 2008.
[2] A. 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.
[3] 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.
[4] 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.
[5] G.L. Danah Boyd and S. Golder, "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter," Proc. 43rd Hawaii
Int'l Conf. System Sciences (HICSS-43), 2010.
[6] A. 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.
[7] P. Galagan, "Twitter as a Learning Tool. Really," ASTD Learning Circuits, p. 13, 2009.
[8] 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.
[9] J. Hightower and G. Borriello, "Location Systems for Ubiquitous Computing," Computer, vol. 34, no. 8, pp. 57-66, 2001.
[10] M. Weiser, "The Computer for the Twenty-First Century," Scientific Am., vol. 265, no. 3, pp. 94-104, 1991.
[11] 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.
[12] 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.
[13] Y. Raimond and S. Abdallah, "The Event Ontology," http:// motools.sf.net/event/event.html, 2007.
[14] T. Joachims, "Text Categorization with Suport Vector Machines: Learning with Many Relevant Features," Proc. 10th European Conf. Machine Learning (ECML '98), pp. 137-142, 1998.
[15] X. Liu, S. Zhang, F. Wei, and M. Zhou, "Recognizing Named Entities in Tweets," Proc. 49th Ann. Meeting of the Assoc. for Computational Linguistics: Human Language Technologies (HLT '11), pp. 359-367, June 2011.