Novateur Publication's
International Journal of Innovation in Engineering, Research and Technology [IJIERT]
ICITDCEME'15 Conference Proceedings
ISSN No - 2394-3696

EARTHQUAKE DETECTION AND REPORTING SYSTEM

Ms. Swapnali Kulkarni
Department of Computer Engineering Pune University VACOE, Ahmednagar, India

Prof. Mandar Kshirsagar
Department of Computer Engineering Pune University VACOE, Ahmednagar, India

ABSTRACT

Now a day's large number of people is moving towards the social networking. Twitter is one of the best examples of social networking. Twitter is generally categorized as a micro-blogging service, which enables users to send brief text updates. These tweets are processed through twitter search API and crawler and used in real time event detection system such as detecting an earthquake. In proposed system semantic analysis is used to analyze the tweets. Support vector machines (SVMs) are used for classification, regression and outliers detection. The occurrence of event detection is done by spatiotemporal model by treating each twitter user as a sensor value. For the purpose of location estimation, an algorithm is used which is based on the tweet content similarity and the location transition history of twitter user.

INTRODUCTION

A social networking service (also called SNS) is a very good platform to build social networks or social relations among people who share their interests, activities or real-life connections. Social networks are web-based services that allow individuals to create a public profile, to create a list of all users with whom we want to share connections, and view and cross the connections within the system. Most social network services are web-based services and allow users to interact over the Internet, such as through e-mail and instant messaging.
Twitter is one of the widely used social networking services that allow you to answer the question,"What are you doing now?. This can be done by sending short text messages to your friends, or"followers." [1]. these short massages are called as tweets. Twitter provides to send the message up to 140 characters. Registered users can read as well as post tweets, but unregistered users can only read tweets. Users can use Twitter via website interface or mobile device app. We generally categorized twitter as a micro blogging service [1]. A difference between traditional blog and micro blog is its contents. Micro blog is typically smaller in both actual and aggregated file size. Generally micro blogs"allow users to share small elements of content such as short sentences, small messages, individual images, or video links". An important characteristic is that, all the micro blogging services are real time in nature. Although blog users typically update their blogs once every several days, Twitter users may write tweets several times per day. Users can know what other users are doing and often what they are thinking about now, users number of times go to the site and check what the other people are doing.
As an application, we develop an earthquake reporting system using tweets. Our system detects an occurrence of earthquake and sends an e-mail, possibly before an earthquake actually arrives at a certain location. An earthquake spreads near about 3 to 7 km per sec. For that reason, a person who is at 100 km apart from an actual earthquake and he is able to communicate and can take the proper action about 20 s before the arrival of an earthquake wave. This paper is organized as described below. In the next section, we explain the proposed system and its structure. After that each module of system is mentioned.

PROPOSED SYSTEM

Proposed system called event detection system detects the target event from twitter. we search from Twitter database and find useful tweets through twitter search API. The flow diagram of this event detection system is shown in figure 1. They have to register their e-mails IDs to receive notifications for future earthquake detection reports. Given system alerts users and urges them to prepare for the imminent earthquake. It is hoped that a user will receive an e-mail notification before the earthquake actually strikes that area.

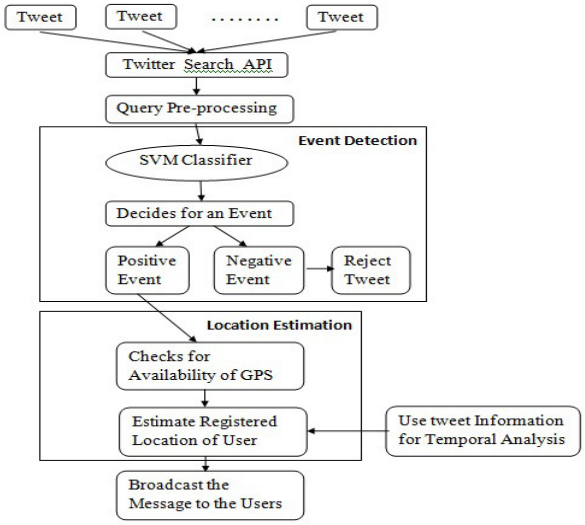

Fig. 1. Method to acquire tweets referred to a target event.

COMPONENTS OF PROPOSED SYSTEM

TWITTER SEARCH API

Twitter search API allows user defined queries against the indices of recent or popular Tweets. IT behaves similarly to, but not exactly like the Search feature available in Twitter mobile or web clients. Search API is basically concentrates on relevance and not completeness. This means that there are chances of missing some tweets or some users from search results. The Search API acts as an interface to the search engine while searching for tweets. In our proposed system using this search API we acquire the tweets having a keyword "earthquake". This tweet retrieval is shown in figure 2.

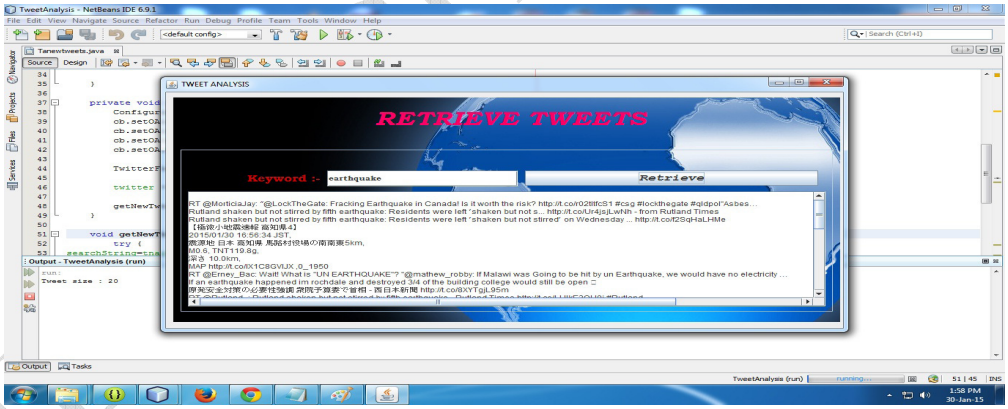

Fig. 2. Tweet retrieval based on keyword earthquake.

QUERY PRE-PROCESSING

Query preprocessing is a technique that prepares the query for optimization purpose. It may change the representation of statement such that the SQL statement Component Integration Services generates is syntactically different from the original statement. Here we are using query pre-processing to extract the tweets more efficiently from twitter database. Hence the query pre-processing technique plays a vital role in the real time system. Following figure 3 represents the GUI of process of results of stop-words removal.

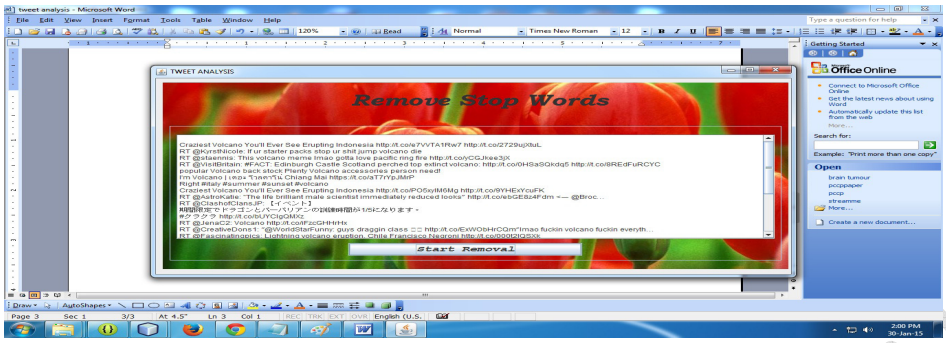

Fig. 3. Stopword removal.

EVENT DETECTION

An event might be anything that happens in the world [2]. Events can be defined as something abnormal, that is, something that happens very rarely in the normal situation. Event detection aims to find such abnormal phenomenon from collected data. We use twitter to detect the target events.

1) Semantic analysis of tweets: Tweets may include information about target event. For example, user might make tweet as Now earthquake is happening or Now it is shaking. Here earthquake and shaking may be a keyword. However user might make the tweet as I am watching movie of earthquake or Someone is shaking hands with my boss. Although the tweet refers to the target event, it may not appropriate as an event report. For instance, Earthquake happened in last year was very harmful or It is an earthquake zone. Such tweets are truly descriptions of the target event, but they are not real-time reports of the events. Therefore it is important to classify the tweets into the classes. One that refers to the actual event occurrence, which we denote as a positive class and remaining as a negative class. We use a support vector machine as a classifier to divide the tweets into positive and negative classes.

Steps for SVM Algorithm:

1) Data preprocessing.

2) Feature Extraction.
   1. Chi- Square
   2. TFIDF
   3. Feature Matrix

3) Performance evaluation.
   1. Precision.
   2. Recall rate.
   3. F-value.

Fig. 4. Steps of SVM Algorithm.

2) Tweet as a sensory value: If user makes the tweet about an earthquake, then it can consider as he is an Earthquake sensor and it returns the positive value. In this paper we have considered two assumptions. Some researchers describe their efforts to extract place names from tweets as a part of Named Entity Recognition [4], [5].

Assumption 1: consider each Twitter user is a sensor. A sensor detects a target event and makes a report about event probabilistically.

Assumption 2: Each tweet is associated with a particular time and location, which is a set of latitude and longitude co-ordinates.

PROBABILISTIC MODEL FOR LOCATION ESTIMATION

Probabilistic model is used for estimating the location of target event.

1) Temporal Model: Each tweet is associated with its post time. We use the temporal model of event detection. The distribution is clearly an exponential distribution. The probability density function of the exponential

distribution is $f(t;\lambda) = \lambda e^{-\lambda.t}$ where t is greater than zero and λ is greater than zero. The exponential distribution occurs naturally when describing the lengths of the inter-arrival times in a homogeneous Poisson process.
2) Spatial Model: Each tweet is having its location. In this paper we describe one method which is used to find the destination of user by using twitter post. In proposed method we choose a base tweet, which is very close to the current users tweet, and then with the help of next set of tweets of base tweet we can predict the destination [6]. For estimating location, given method uses following functions.
we define the similarity between user us current tweet ui and the base tweet cj as
Similarity ui;cj = α.PastSimilarity u;c + (1 - α).Current Similarity ui;cj
Where, α is a weighting parameter.
Past similarity is calculated with the help of tweet similarity and location similarity between user u and base user c. PastSimilarityu;c = β.TweetSimilarityu;c + (1 - β).LocationSimilarityu;c
Where, β is a weighting parameter.

EARTHQUAKE REPORTING SYSTEM

We developed an earthquake-reporting system using the event detection algorithm. Earthquake information is very much valuable if it is received in real time. Given some amount of advanced warning, any person would be able to turn off a domestically used gas or heater at home and then find protection under a desk or table if such a person were to have some seconds notification before an earthquake actually strikes an area. Large amount of work is done on seismology. Also much more efforts have been taken to produce short-term forecasts to realize an earthquake warning system by observing electromagnetic emissions from ground-based sensors and satellites.

CONCLUSION

In this paper we have studied, the real time nature of twitter and used it in the event detection system. Semantic analysis is applied to the tweets to classify the tweets into the positive and negative class. Semantic analysis is used for detecting the target events. We consider each twitter user as a sensor for particular event, and set that problem as detection of an event based on sensory observations. For location estimation a method is proposed based on tweet content similarity and the past similarity of tweet. As an application, we developed an earthquake reporting system, which is a best approach to notify people promptly of an earthquake event.

REFERENCES

[1] M. M. Sarah, C. Abdur, H. Gregor, L. Ben, and M. Roger, Twitter and the Micro Messaging Revolution, technical report, OReilly Radar, 2008.
[2] Y. Y. Raimond and S. Abdallah, The Event Ontology, http:// motools. sf.net/event/event.html, 2007.
[3] T. 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.
[4] X. 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.
[5] A. A. Ritter, S. Clark Mausam, and O. Etzioni, Named Entity Recognition in Tweets: An Experimental Study, Proc. Conf. Empirical Methods in Natural Language Processing, 2011.
[6] T. Takuya Shinmura, Dandan Zhu, Jun Ota, Yusuke Fukazawa Destination Prediction Considering Both Tweet Contents and Location Transition Hitstory, ICMU, pp95-96, 2014.