I J C T A, 9(19) 2016, pp. 9101-9105

Identifying the Event type and Event Location using Sentiment Analysis on Tweets

Vijay Singh*, Bhasker Pant* and Devesh Pratap Singh*

ABSTRACT

Social Media is one of the fastest medium of information broadcasting. In this research, this inherent property of the social network and apply machine learning algorithm to identify the specific event and the location where the event occurred. This leads to the quick response by the concerned authorities and saves the lives of many people in case of emergency. Events can be earthquakes Location Based Services(LBS), terrorist attack, riots, cloudburst, major traffic congestion, an epidemic, etc. By experimental results of this proposed system, it is noticed that, the method can maintain a pretty good performance.

Keywords: Sentiment analysis, machine learning, event detection, location Based Services(LBS).

I. INTRODUCTION

One of the most powerful media is social networks in today's digital world. In the current scenario Social Network services plays an important role in everyone life. Twitter is one of the effective micro-blogging services started in July 2006, Where the user can send and read short messages of 140 character length. There are around 40% tweets are conversational tweets and the rest are pointless babbles, news, spams, etc. There are various applications have been developed using these conversational tweets and analyze it at many dimensions, like event detection, location detection, traffic jam prediction,anomaly detection. Jalal Mahmud at el. Developed an algorithm to predict the home location of the user by their twitting pattern and by analysis their twitting contents [1]. Some previous studies show that due to information transmission delay situation of the event location become critical. It can be avoided if the event detected instantly and based on the severity of the event , further pass the message to the concerned authorities. It works like an alarming system in case of the specific event or any disastrous condition. Yusuke Hara analyzed twitter data at the time of Great East Japan earthquake. He used behavior inference system and geotag data for prediction [2]. Ji Ao at el. contributed how location can be determined by combining content-based location,posting location and registering the location of the user. Their results show significant improvement in location estimation [3]. Chenliang Li and Aixin Sun, implementsFine-Grained location extraction from tweets, by combining lexical features, grammatical features and geographical feature [4]. Traditionally event were detected by TF-IDF algorithm to track and detect specific contents. The main contribution of this paper is, it successfully detect event and their event location by a very effective and simple procedure. Zhicong Tan et al. used multi-layer event detection algorithm to detect local hot event and global hot events and shows that, how much effieient in event cluster generation [7]. P. G. et al. proposed a technique by using support vector machine for tweets classification and causal relation is used for detect the relation between two events. causality is a term used to define the relation between two events [8]. Xu. f et al [9] used a bootstrapping technique for detecting events and event type automatecially. McCraker , N et al [10]

* Department of Computer Science and Engineering, Graphic Era University, Dehradun, Uttarakhand, India, *E-mails: vijaysingh_agra@hotmail. com; pantbhaskar2@gmail. com; devesh. geu@gmail. com*

used a hybrid approach of knowledge based technique and statistical technique an event. Duc T. Nguyen et al [11], proposed a technique of early event identification. They did behavioral analysis on social network big data scenario, and their system works on real time. Mariam Adedoyin-Olowe et al [12] proposed a technique of Transaction based rule change technique to extract content about a particular event. They used hashtag keywords and Association Rule mining for detect specific topic.

The rest of the work is organized as follows. The proposed algorithm for event detection and event location detection are explained in section II. Experimental results are presented in section III. Concluding remarks are given in section IV.

II. PROPOSED ALGORITHM

2.1. Event detection algorithm

Many users are using social media for updating real life event. These events can be classified as general events and specific event or alarming event. General events can be ignored, but specific can't because users may be required some kind of help or they want to inform other users about to come emergency situation. various studies have been done in this domain to detect the specific event in routine conversation of the users. Jianshu Weng and Bu-Sung Lee proposed a system in which they detect event by simply content analysis with clustering of wavelet-based signals [5]. Zhicong Tan et al, develop a multi-layer system for event detection. In which top tweets are collected in a particular group for the reliabililty of the system and formulate event cluster to group members [6].

In this section event detection algorithm is proposed and the steps involved shown in figure 1. This section is organized as follows:

Generate Event Model

- Collection of tweets
- Noise Removal
- Stop word removal
- Calculate Term Frequency
- Create event model

Figure 1: Steps for generating Event Model

Different event model are generated by the collection of tweets. In this work, we collected 20,000 event specific tweets, like 20,000 tweets representing terrorist attack tweets, 20,000 earthquake oriented tweets and so on. Then apply basic noise removal techniques and remove stopwords for performance improvement. Then calculate the frequency of remaing tokens generated.

$$E_1, E_2, E_3 \ldots\ldots E_n$$

Where E is a SET of event specific tokens, like terrorist attack, traffic congestion,cloudburst,epidemic etc.

In the similar ways normal days tweets are calculated (G). Then apply E difference G to generate more specific tokens about the event, shown in figure 2.

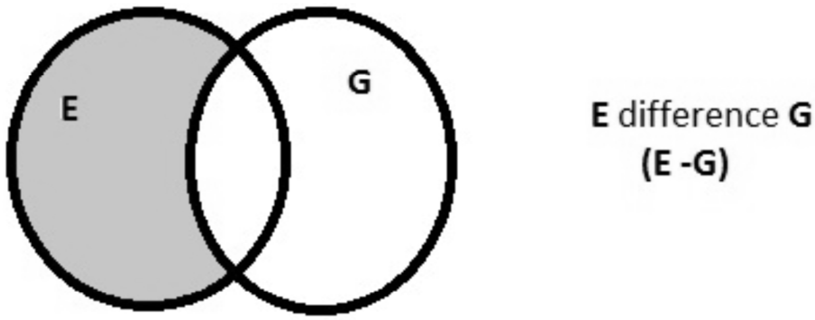

Figure 2: Venn diagram between specific and general event

$$S_e=E_i-G \quad \text{eq (1)}$$

Where Se is Specific event. Now download the real time tweets at different intervals and apply the same techniques mention is subsection I, and map with the existing model. If the collected tweets tokens map with the existing token more than 30% then the event occour otherwise consider as routine conversation. mapping is shown in figure 3.

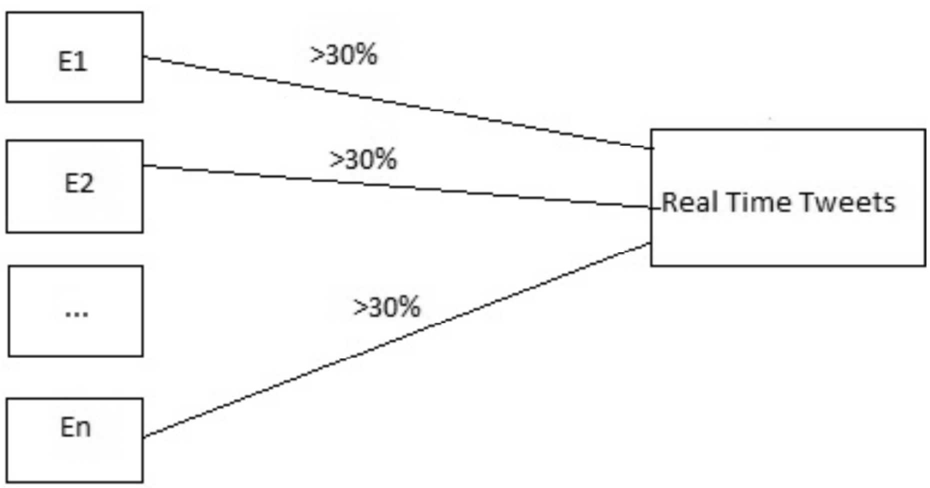

Figure 3: Event detection

Algorithm: Event Model Generation

//INPUT*: Different category of tweets(E),General tweets(G)*

//OUTPUT: *Event Detection*

Specific_token()

For each category i

Token[i]<- *generate_token(*E_i*)*

FREQ[i]<-*token_frequency(*E_f*)*

Specific_token[i]<- *set_difference(Ef,G)*

If(specific_token(i)>30%) then event occour

Otherwise repeat the steps

2.2. Event location detection algorithm

Second objective of the paper is how to locate the location where the specific event occour. various have been publish in this domain , most of them uses content based and geotag based approach for event location detection. In this reasech we use google fusion table to identify the event location. The output of the event detection algorithm is the input of google fusion table in the CSV format. We appalied this procedure for terrorist and earthquake specific events and find the accurate result, shown in figure 3 and figure 4 respectively.

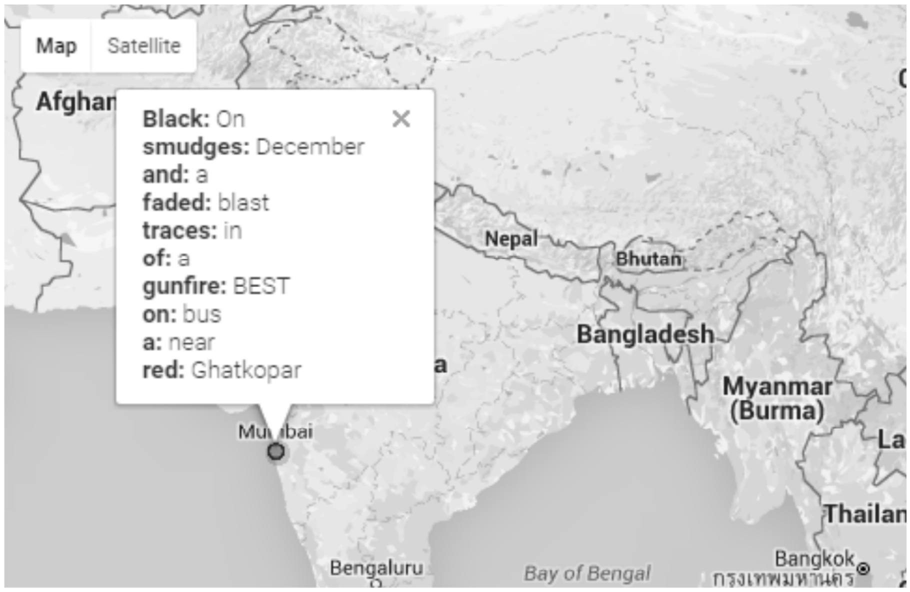

Figure 3: Terrorist attack location on the Map

Figure 4: Earthquake location on the Map

III. EXPERIMENT AND RESULT

For the demonstration of the above mentioned methodology real time tweets are used for model generation and for event location detection. For noise and stopword removal “tm” package of “R” is used. The PC for experiment is equipped with an Intel P4 2. 4GHz Personal computer and 4GB RAM. The proposed scheme is tested is applied on various events and find the system accurate more than 67 %.

IV. CONCLUSION

In this paper we have shown, event and event location can be detected by simple sentiment analysis and machine learning algorithm. Using the google fusion table, we achive a good performance. The significance behind this proposed work, it uses very simple procedure for event detection and event type detection and it correctly achived the desired task. In the future work we will try to develop self-generated event model so the performance of the system can be improved.

REFERENCES

[1] Jalal Mahmud, Jeffrey Nicholos,Clemens rews "Home Location Identification of Twitter users", ACM Transations on Intelligent Systems and Technology, Vol. %, No. ,Article 47, July 2014.

[2] Yusuke Harai, "Behaviour analysis using tweet data and geo-tag data in natural disaster," Transportation Research Procedia 11(2015)399-412.

[3] Ji Ao,Peng Zhang,Yanan Cao, "Estimating the location of Emergency Event from Twitter Streams," 2nd International Conference on Information Technology and quantitative Management, ITQM, Procedia Computer Science 31(2014)731-739.

[4] Chenliang Li and ixin Sun, "Fine-grained location extraction from tweets with temporal awareness", SIGIR'14 Proceedings of the 37th international ACM SIGIR conference on Research and development in information retrieval pages 43-52.

[5] Jianshu Weng and Bu-Sung Lee, "Event detection in Twitter", Proceedings of the fifth International AAI Conference on Weblogs and Social Media (2011).

[6] Zhicong Tan,Peng Zhang, Jianlong Tan and I Guo,"A Multi-layer Eevnt Detection Algorithm for Detecting Global and Local Hot vents in Social Networks",ICCS 2014, 14th International Conference on Computational Science,Procedia Computer Science, Volume 29, 2014, Pages 2080-2089.

[7] Zhicong Tan, Peng Zhang, Jianlong Tan and li Guo, "A Multi-layer Event Detection Algorithm for Detecting Global and Local Hot Events in Social Networks", ICCS 2014, 14th International Conference on Computational Science, Procedia Computer Science, Volume 29, 2014, Pages 2080-2089.

[8] P. G Preethi, V. Uma and Ajit Kumar,"Temporal Sentiment Analysis and Causal Rules Extraction from Tweets for Event Prediction",International Conference on Intelligent Computing, Communication & Convergenc (ICCC-2015),Procedia Computer Science 48(2015)84-89.

[9] Xu, F. , Uszkoreit, H. , Li, H. 2006. Automatic Event and Relation Detection with Seeds of Varying Complexity, In Proceedings of the AAAI 2006 Workshop Event Extraction and Synthesis, Boston, 491-498.

[10] Naughton, M. Stokes, M. and Carthy, J. 2008. Investigating Statistical Techniques for sentence level event classification, In *Proceedings of the 22nd International Conference on Computational Linguistics*, 617-624.

[11] Duc T. Nguyen, Jai E. Jung, "Real-time event detection for online behavioral analysis of big social data", Future Generation Computer System,2016.

[12] Mariam Adedoyin-Olowe, Mohamed Medhat Gaber, Carlos M. Dancausa, Frederic Stahl and Joao Bartolo Gomes, "A rule dynamics approach to event detection in Twitter with its application to sports and politics" Expert Systems with Applications, 15 August 2016, Pages 351-360.