RESEARCH ARTICLE OPEN ACCESS

Real-Time Detection of Traffic from Social Media's Stream or Post's Analysis.

Ramiz Shaikh[1], Sameer Shinde[2], Vaibhav Pawar[3]
[1,2,3](Department of Computer Engineering, L.G.N Sapkal college of Engg, Anjneri hill, Trambakeshwer, Nashik)

Abstract:

Social networking sites are source of information for event detection, with specific reference of the road traffic activity blockage and accidents or earth-quack sensing system. During this paper, we have a tendency to present a time period observation system supposed for traffic occasion detection coming back from social media stream analysis. The system fetches tweets coming back from social media/network as per a many search criteria; ways tweets/posts, by applying matter content mining methods; last however not least works the classification of social networks posts. The goal is to assign appropriate category packaging to each posts, as a result of connected with Associate in Nursing activity of traffic event or maybe not. The traffic recognition system or framework was utilized for time period observation of varied areas of the road network, taking into consideration detection of traffic occasions simply virtually in actual time, frequently before on-line traffic news sites. All people utilized the support vector machine sort of a classification unit; what is more, we tend to accomplish a good accuracy price of 95.76% by making an attempt a binary classification problem. All people were conjointly capable to discriminate if traffic is triggered by Associate in nursing external celebration or not, by partitioning a multi category classification issue and getting accuracy price of 88.89.

Keywords — **Social media, Traffic detection, Text mining; Privacy, Service oriented architecture (SOA), machine learning, Twitter/social media's stream analysis.**

I. INTRODUCTION

Social media platforms square measure wide used for distributed data regarding the detection of events, like traffic block, incidents, natural disasters (earthquakes, storms, fires, etc.), or alternative events. An occasion is outlined as a true world existence that happens in an exceedingly definite time and house. Usually traffic connected events; individuals oftentimes share by suggests that of add data regarding this traffic state of affairs around them whereas driving. For this purpose, event detection from social networks is additionally usually utilized with Intelligent Transportation Systems (ITSs). ITSs afford, e.g., period data regarding weather, tie up or regulation, or set up economical (e.g., shortest, quick driving, least polluting) routes. Event detection from social networks investigation could be an additional stimulating downside than event detection from Ancient broadcasting like blogs, emails, etc. In fact, SUMs square measure unstructured and unequal texts; it holds informal or shortened words, mistakes or grammatical errors. SUMs contain a large quantity of not helpful or unarticulated data that has got to be processed. In step with Pear Analytics, it's been calculable that over forty first of all posts SUMs (i.e., tweets) are not sensible with any helpful information for the audience. For all of those reasons, so as to research the info coming back from social networks or text mining techniques, we have a tendency to use to extract vital information, of information mining, device learning, numbers, and tongue process (NLP). During this paper, we have a tendency to propose a superb system, based mostly upon text mining and instrumentality learning algorithms, for current detection of traffic occasions from social media stream analysis. The system, once having a feasibleness study, offers been designed and created from the bottom as a result of an event-driven infrastructure, created on a Service targeted design (SOA). The program exploits offered technologies

established on state-of the-art processes for text message analysis and pattern class. These technologies and strategies are analysed, configured, adapted, and integrated to be able to build the intelligent system. specifically, we have a tendency to gift an excellent experimental study, that options been performed for deciding the foremost effective between totally different state-of the-art approaches supposed for text classification. The chosen picked approach was integrated in to the last system and used for the on-the field current detection of traffic things.

II. OBJECTIVES

1. Design a real-time detection system for traffic analysis.
2. The vital objectives of proposed system is as follows:- Design a real-time detection system for traffic analysis.
3. The aim is to assign suitable class label to every tweet, as related with an activity of traffic event or not.
4. It performs a multi-class classification, which recognizes non-traffic, traffic due to congestion or crash, and traffic due to external events.
5. It detects the traffic events in real-time and It is developed as an event-ambitious infrastructure, built on an SOA architecture.

III. RELATED WORK

Text mining means to the process of automatic extraction of important information and knowledge from unstructured text. Text mining is a difference on a field called data mining that tries to find interesting pattern from large databases. Text mining, also known as Rational Text Exploration, Text Data Mining or Knowledge-Discovery in Text (KDT), refers mostly to the process of extracting interesting and non-trivial information and knowledge from unstructured text. As most data (over 80%) is stored as text, text mining is supposed to have a high commercial potential value. Knowledge may be exposed from many sources of information, yet, formless texts remain the largest readily existing source of knowledge. Most of text mining methods are based on the idea that a document can be faithfully represented by the set of words contained in it (bag-of-words representation).Data mining and device learning algorithms (i.e., support vector machines (SVMs), decision trees, neural networks, etc.) are applied to the documents in the vector space representation, to build classification, clustering or regression models. In this way we aim to support traffic and city administrations for managing scheduled or unexpected events in the city.

Now days the social media is used for the extraction of information related to the real time event detection. Small events have a small number of SUMs related to them [1], belong to a precise geographic location, and are concentrated in a small time interval [9]. Large events such as earthquakes, tornados are detected by a large number of SUMs, and by a wider geographic coverage [9].

Sakaki et al. [6] used twitter streams to detect earthquakes etc. by monitoring special trigger-keywords, and by applying an SVM as a binary classifier of positive events and negative events. In [7], authors presented a method for detecting real world events, such as natural disasters, tornados by analysing Tweets by employing both term-frequency-based techniques and NLP [7]. In paper [2] authors proposed a novel system which detect and analyse events from rich information from Twitter. The proposed method supports the following three functionalities (1) detecting new events, (2) ranking events according to their importance, and (3) generating temporal and spatial patterns for events [2]. Initially authors focused on Crime and Disaster related Events (CDE), such as shooting, car accidents etc. [2].

Vikram Singh et.al. proposed an effective tokenization technique which is based on training. In the results, it is shown that tokenization along with pre-processing generates better tokens [4], If less number of token generated then less storage space is required and it facilitates more accuracy in results retrieval [4]. Algorithm also responsible for reducing the time of information retrieval model [4].

Maximilian Walther et. al. in [5] detect Geo-spatial Event using twitter. The paper describes a new scenario and approach to tackle it. In this approach they gathered tweets for target events that can be defined by a user via keywords [5]. Classification and particle filtering methods are used for detecting events. Authors used common theme as if people tweeting from the same place use the same words which means that they talking about that thing only [5].

IV. EXISTING SYSTEM

Existing system propose an intelligent system, based on text mining and machine learning algorithms, for real-time detection of traffic events from Twitter stream analysis. The system, after a feasibility study, has been designed and developed from the ground as an event-driven infrastructure, built on a Service Oriented Architecture (SOA) [1]. The system exploits available technologies based on state-of the-art techniques for text analysis and pattern classification [4]. These technologies and techniques have been analysed, adapted, and added with existing in order to build the intelligent system [1]. In particular, system present an experimental study, which has been performed for determining the most effective among different state-of-the art approaches for text classification. The chosen approach is added into the final system and then used for the on-the-field real-time detection of traffic events.

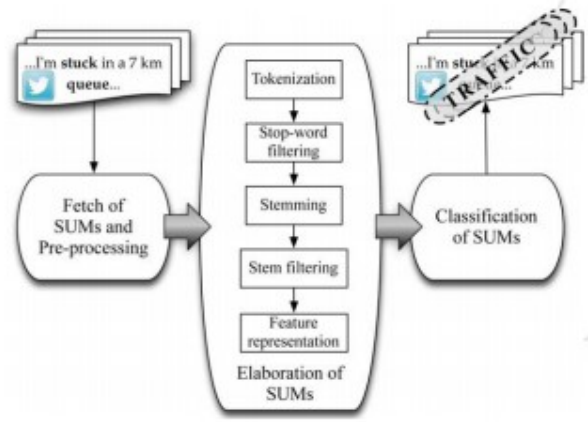

Figure 1: System architecture for traffic detection from Twitter stream analysis.

V. PROPOSED WORK

We propose associate degree intelligent system, supported text mining and machine learning algorithms, for time period detection of traffic events from Twitter stream analysis. The system, when a practicability study, has been designed associate degrade developed from the bottom as an event-driven infrastructure, designed on a Service homeward design (SOA). The system exploits on the market technologies supported progressive techniques for text analysis and pattern classification. These technologies and techniques are analysed, tuned, adapted, and integrated so as to make the intelligent system. Above all, we have a tendency to gift associate degree experimental study, that has been performed for decisive the foremost effective among completely different progressive approaches for text classification. The chosen approach was integrated into the ultimate system and used for the on-the-field time period detection of traffic events.

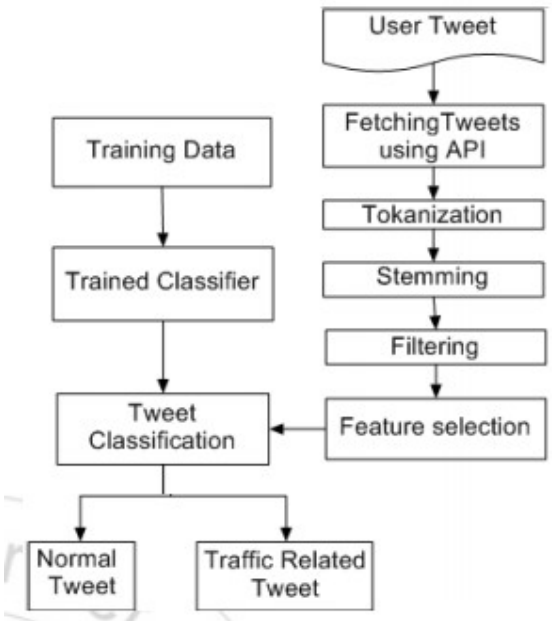

Figure 2: Proposed Architecture

VI. MATHEMATICAL MODEL

Let S is the Whole System Consists:

S= {I, P, O}
I = Input.
P= Procedure.
O= Output

I = {U, T, TS, url, Tk}.

1. Let U is set of number of twitter users in the system. U = {u1, u2,……un}.
2. T is set of number Tweet or status update of twitter user. T= {t1, t2, t3…tn}.
3. TS is twitter streamer who analyses the twits.
4. url is the URL of twitter user who have updated status.
5. Tk is the tokenization of SUM where, SUM is the Status Update Message of twitter user.

VII. CONCLUSIONS

In this work, we've projected a system for period detection of traffic-related events from Twitter stream analysis. System is in a position to fetch and classify streams of tweets and to apprise the users of the presence of traffic events. We advancing explored the authentication yet as trust and name calculation and management of CSPs and SNPs, that square measure 2 terribly vital and hardly explored problems with reference to CC and WSNs integration. Further, we tend to project a unique ATRCM system for CC-WSN integration. Discussion and analysis concerning the authentication of CSP and SNP yet because the trust and name with reference to the service provided by CSP and SNP are bestowed followed with elaborate style and practicality analysis concerning the projected ATRCM system. Of these incontestable that the projected ATRCM system achieves the subsequent 3 functions for CC-WSN integration:

1) Supportive CSP and SNP to avoid malicious impersonation attacks.
2) Shrewd and handling trust and name concerning the service of CSP and SNP.
3) Serving to CSU select needed CSP and helping CSP in choosing acceptable SNP,

Based on

(i) The credibleness of CSP and SNP;
(ii) The attribute demand of CSU and CSP;
(iii) The value, trust and name of the service of CSP and SNP.

ACKNOWLEGMENT

It provides us great pleasure in presenting the preliminary project report on Real-Time Detection of Traffic From Social Media's Stream Analysis. I would like to take this opportunity to thank my internal guide Prof. P. A. Kale and project coordinator Prof. Nandwalkar, for giving me all the support and guidance I needed. I am really grateful to them for their kind support. Their valuable ideas were very helpful. I am also grateful to Prof. N. R. Wankhade, Head of Computer Engineering Department, L. G. N. Sapkal Collage of Engineering for his essential support, suggestions. In the end our special thanks to College Management and all Staff for providing several resources such as laboratory with all needed software platforms, continuous Internet connection, for Our Project.

REFERENCES

1. AminaMadani, Omar Boussaid, Djamel Eddine Zegour and Algiers, Algeria, "What's Happening: A Survey of Tweets Event Detection," J. Internet Services Appl., vol. 1, no. 1, pp. 7–18, 2010.
2. Parikh, Kamalakar Karlapalem., "ET: Events from Tweets," Future Generat. Comput. Syst., vol. 25, no. 6, pp. 599–616, Jun. 2009. [2]
3. Alan Mislove, MassimilianoMarcon, Krishna P. Gummadi, Peter Druschel, Bobby Bhattacharjee., "Measurement and Analysis of Online Social Networks," Proc. IEEE, vol. 99, no. 1, pp. 149–167, Jan. 2011. [3]
4. B. Chen and H. H. Cheng, "A review of the applications of agent technology in traffic and transportation systems," IEEE Trans. Intell. Transp. Syst., vol. 11, no. 2, pp. 485–497, Jun. 2010. [4]

5. T. Sakaki, Y. Matsuo, T. Yanagihara, N. P. Chandrasiri, and K. Nawa, "Real-time event extraction for driving Programming, Bangalore, India, Jan 2010, pp. 147-158. [5]
6. M. Yuriyama and T. Kushida, "Sensor-cloud infrastructure Physical sensor management with virtualized sensors on cloud computing," in Proc. 13th Int. Conf. Netw.,-Based Inf. Syst., Sep. 2010, pp. 1–8 [6]
7. Chen and Harry H., "A Review of the Applications of Agent Technology in Traffic and Transportation Systems," Wireless Commun. Mobile Comput, vol. 14, no. 1, pp. 19–36, Jan. 2014. [7]
8. Álvaro González, Luis M. Bergasa and J. Javier Yebes, "Chen and Harry H.," ACM Trans. Sensor Netw., vol. 5, no. 2, Mar. 2009, Art. ID 10. (2002) The IEEE website. [Online]. Available: http://www.ieee.org/
9. Vikram Singh and Balwinder Saini ―An Effective Tokenization Algorithm for Information Retrieval System‖ CS & IT-CSCP 2014 [6]
10. Maximilian Walther and Michael Kaisser, ‖Geo-spatial Event Detection in the Twitter Stream‖, P. Serdyukov et al. (Eds.): ECIR 2013, LNCS 7814, pp. 356367, 2013.cS pringer VerlagBerlinHeidelberg 2013.

AUTHORS

Ramiz Khalik Shaikh pursuing BE in *Department Of Computer Engineering, L.G.N Sapkal college of Engg, Anjneri hill, Trambakeshwer, Nashik.*

Sameer Balasaheb Shinde pursuing BE in *Department Of Computer Engineering, L.G.N Sapkal college of Engg, Anjneri hill, Trambakeshwer, Nashik.*

Vaibhav Natha Pawar pursuing BE in *Department Of Computer Engineering, L.G.N Sapkal college of Engg, Anjneri hill, Trambakeshwer, Nashik.*