International Journal of Research
Available at https://edupediapublications.org/journals

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 10
September 2017

Apprehensive Of Transportation Through Twitter Analysis

Vinay Tentu & Dr. K Rajendra Prasad Ph.D
M.Tech (Software Engineering) Institute Of Aeronautical Engineering, Hyd
Professor And Head Of Cse Institute Of Aeronautical Engineering, Hyd

Abstract:

Internet sites are source of info for event detection, with specific mention of the road traffic activity blockage and accidents or earth-quack sensing system. In this paper, we present a real-time monitoring system intended for traffic occasion detection coming from Twitter stream analysis. The system fetches tweets coming from Twitter as per a several search criteria; methods tweets, by applying textual content mining methods; last but not least works the classification of twitter posts. The goal is to assign suitable class packaging to every tweet, because related with an activity of traffic event or perhaps not. The traffic recognition system or framework was utilized for real-time monitoring of various areas of the street network, taking into account detection of traffic occasions just almost in actual time, regularly before on-line traffic news sites. All of us employed the support vector machine like a classification unit; furthermore, we accomplished a great accuracy value of ninety five. 75% by attempting a binary classification issue. All of us were also capable to discriminate if traffic is triggered by an external celebration or not, by resolving a multiclass classification issue and obtaining accuracy worth of 88. 89%.

Keywords: Traffic Event Detection, Tweet Classification, Text Mining, And Social Sensing.

I. INTRODUCTION

Twitter is prone to malicious tweets containing URLs for spam, phishing, and malware distribution. Conventional Twitter spam detection schemes utilize account of features such as the ratio of tweets containing URLs and the account creation date, or relation features in the Twitter graph. These detection schemes are ineffective against feature fabrications or consume much time and resources. Conventional suspicious URL detection schemes utilize several features including lexical features of URLs, URL redirection, HTML content, and dynamic behavior. However, evading techniques such as time-based evasion and crawler evasion exist. In this paper, we 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). The system exploits available technologies based on state-of-the-art techniques for text analysis and pattern classification. These technologies and techniques have been analyzed, tuned, adapted, and integrated in order to build the intelligent system. In particular, we 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 was integrated into the final system and used for the on-the-field real-time detection of traffic events.

In the existing system attackers use shortened malicious URLs that redirect Twitter users to external attack servers. To cope with malicious tweets, several Twitter spam detection schemes have been proposed. These schemes can be classified into account feature-based, relation feature-based, and message feature based schemes. Account feature-based schemes use the distinguishing features of spam accounts such as the ratio of tweets containing URLs, the account creation date, and the number of followers and friends. However, malicious users can easily fabricate these account features. The relation feature-based schemes rely on more robust features that malicious users cannot easily fabricate such as the distance and connectivity apparent in the Twitter graph. Extracting these relation features from a Twitter graph, however, requires a significant amount of time and resources as a Twitter graph is tremendous in size. The message feature-based scheme focused on the lexical features of messages. However, spammers can easily change the shape of their messages. A number of suspicious URL detection schemes have also been introduced. With reference to current approaches for using social media to extract useful information for event detection, we need to distinguish between small-scale events and large-scale events. Small-scale events (e.g., traffic, car crashes, fires, or local manifestations) usually have a small number of SUMs related to them, belong to a precise geographic location, and are concentrated in a small time interval.

On the other hand, large scale events (e.g., earthquakes, tornados, or the election of a president) are characterized by a huge number of SUMs, and by a wider temporal and geographic coverage. Consequently, due to the smaller number of SUMs related to small-scale events, small-scale event detection is a non-trivial task. Several works in the literature deal with event detection from social networks. Many works deal with large-scale event detection, and only a few works focus on small-scale event. Regarding small-scale event detection, the detection of fires in a factory from Twitter stream analysis, by using standard NLP techniques and a Naive Bayes (NB) classifier in this project, we focus on a particular small-scale event, i.e., road traffic, and we aim to detect and analyze traffic events by processing users' SUMs belonging to a certain area and written in the Italian language. To this aim, we propose a system able to fetch, elaborate, and classify SUMs as related to a road traffic event or not.

II. EXISTING AND PROPOSED SYSTEMS

A. Existing System

Recently, social networks and media platforms have been widely used as a source of information for the detection of events, such as traffic congestion, incidents, natural disasters (earthquakes, storms, fires, etc.), or other events. Sakaki et al. use Twitter streams to detect earthquakes and typhoons, by monitoring special trigger-keywords, and by applying an SVM as a binary classifier of positive events (earthquakes and typhoons) and negative events (non-events or other events). Agarwal et al. focus on the detection of fires in a factory from Twitter stream analysis, by using standard NLP techniques and a Naive Bayes (NB) classifier. Li et al. propose a system, called TEDAS, to retrieve incident-related tweets. The system focuses on Crime and Disaster-related Events (CDE) such as shootings, thunderstorms, and car accidents, and aims to classify tweets as CDE events by exploiting a filtering based on keywords, spatial and temporal information, number of followers of the user, number of rewets, hash tags, links, and mentions.

Disadvantages of Existing System:

- Event detection from social networks analysis is a more challenging problem than event detection from traditional media like blogs, emails, etc., where texts are well formatted.
- SUMs are unstructured and irregular texts; they contain informal or abbreviated words, misspellings or grammatical errors.
- SUMs contain a huge amount of not useful or meaningless information

B. Proposed System

In this paper, we 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). The system exploits available technologies based on state-of-the-art techniques for text analysis and pattern classification. These technologies and techniques have been analyzed, tuned, adapted, and integrated in order to build the intelligent system as shown in Fig.1. In particular, we 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 was integrated into the final system and used for the on-the-field real-time detection of traffic events. In this paper, we focus on a particular small-scale event, i.e., road traffic, and we aim to detect and analyze traffic events by processing users' SUMs belonging to a certain area and written in the Italian language. To this aim, we propose a system able to fetch, elaborate, and classify SUMs as related to a road traffic event or not. To the best of our knowledge, few papers have been proposed for traffic detection using Twitter stream analysis. However, with respect to our work, all of them focus on languages different from Italian, employ different input features and/or feature selection algorithms, and consider only binary classifications.

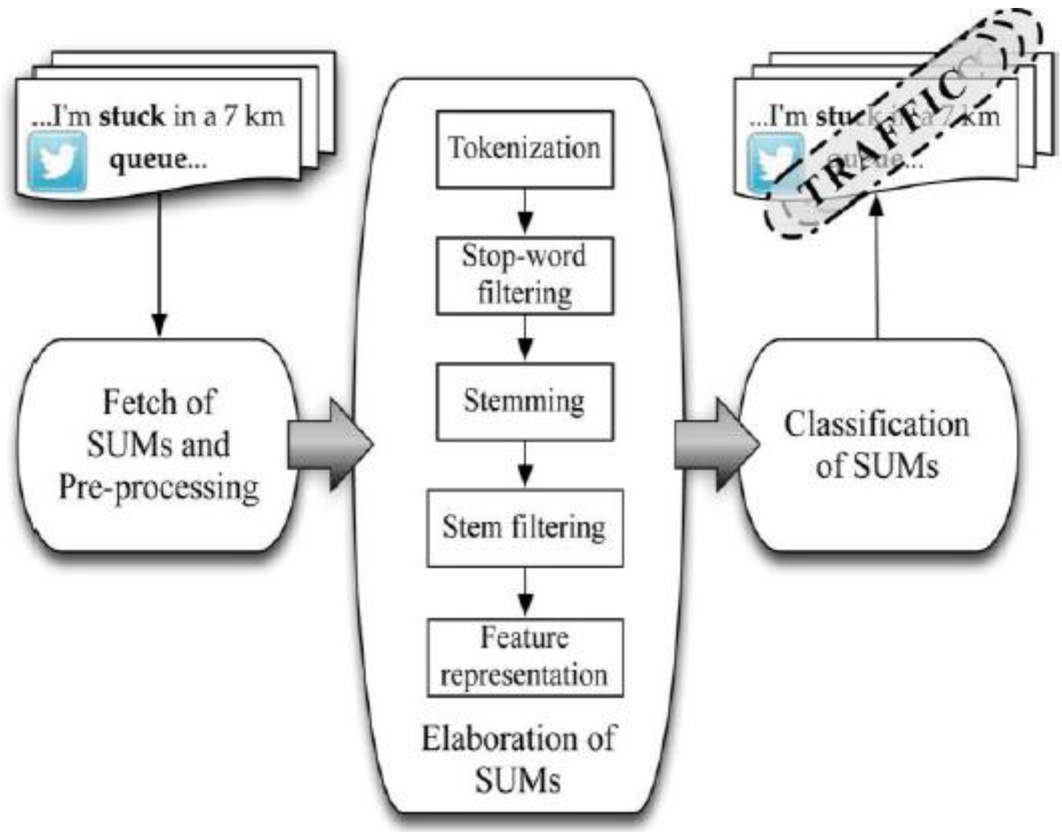

Fig.1. System Architecture.

Advantages of Proposed System:

1. Tweets are up to 140 characters, enhancing the real-time and news-oriented nature of the platform. In fact, the life-time of tweets is usually very short, thus Twitter is the social network platform that is best suited to study SUMs related to real-time events.
2. Each tweet can be directly associated with meta-information that constitutes additional information.
3. Twitter messages are public, i.e., they are directly available with no privacy limitations. For all of these reasons, Twitter is a good source of information for real-time event detection and analysis.
4. Moreover, the proposed system could work together with other traffic sensors (e.g., loop detectors, cameras, infrared cameras) and ITS monitoring systems for the detection of traffic difficulties, providing a low-cost wide coverage of the road network, especially in those areas (e.g., urban and suburban) where traditional traffic sensors are missing.
5. It performs a multi-class classification,which recognizes non-traffic, traffic due to congestion or crash, and traffic due to external events
6. It detects the traffic events in real-time; and iii) it is developed as an event-driven infrastructure, built on SOA architecture.

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 10
September 2017

III. LITERATURE REVIEW

A. What's happening: A Survey of Tweets Event Detection?

Twitter is now one of the main means for spread of ideas and information throughout the Web. Tweets discuss different trends, ideas, events, and so on. This gave rise to an increasing interest in analyzing tweets by the data mining community. Twitter is, in nature, a good resource for detecting events in real-time. In this survey paper, authors have presented four challenges of tweets event detection: health epidemics identification, natural events detection, trending topics detection, and sentiment analysis. These challenges are based mainly on clustering and classification. We review these approaches by providing a description of each one. These last years have been marked by the emergence of micro blogs. Their rates of activity reached some levels without precedent. Hundreds of millions of users are registered in these micro blogs as Twitter. They exchange and tell their last thoughts, moods or activities by tweets in some words.

B. ET: Events from Tweets

Social media sites such as Twitter and Face book have emerged as popular tools for people to express their opinions on various topics. The large amount of data provided by these media is extremely valuable for mining trending topics and events. In this paper, we build an efficient, scalable system to detect events from tweets (ET). Our approach detects events by exploring their textual and temporal components. ET does not require any target entity or domain knowledge to be specified; it automatically detects events from a set of tweets. The key components of ET are: (1) an extraction scheme for event representative keywords (2) an efficient storage mechanism to store their appearance pat- terns, and (3) a hierarchical clustering technique based on the common co-occurring features of keywords. Authors presented a scalable and efficient system, called ET, to detect real world events from a set of micro blogs/tweets. The key feature of this system is the efficient use of con- tent similarity and appearance similarity among keywords, to cluster the related keywords. We demonstrate the effectiveness of this combination in our experiments. ET does not need any human expertise or knowledge from other sources like Wikipedia, and still provides very accurate results. ET is evaluated on two different datasets from two different domains and it yields great results for both of them in terms of the precision.

C. Measurement and Analysis of Online Social Networks

Online social networking sites like Orkut, YouTube, and Flickr are among the most popular sites on the Internet. Users of these sites form a social network, which provides a powerful means of sharing, organizing, and finding content and contacts. The popularity of these sites provides an opportunity to study the characteristics of online social network graphs at large scale. Understanding these graphs is important, both to improve current systems and to design new applications of online social networks. This paper presents a large-scale measurement study and analysis of the structure of multiple online social networks. We examine data gathered from four popular online social networks: Flickr, YouTube, Live Journal, and Orkut. We crawled the publicly accessible user links on each site, obtaining a large portion of each social network's graph. Our data set contains over 11.3 million users and 328 million links. We believe that this is the first study to examine multiple online social networks at scale. Our results confirm the power-law, small-world, and scale free properties of online social networks. We observe that the in degree of user nodes tends to match the out degree; that the networks contain a densely connected core of high-degree nodes; and that this core links small groups of strongly clustered, low-degree nodes at the fringes of the network.

Finally, the implications of these structural properties for the design of social network based systems. Presented an analysis of the structural properties of online social networks using data sets collected from four popular sites our data shows that social networks are structurally different from previously studied networks, in particular the Web. Social networks have a much higher fraction of symmetric links and also exhibit much higher levels of local clustering. We have outlined how these properties may affect algorithms and applications designed for social networks.

D. Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors

Twitter, a popular micro blogging service, has received much attention recently. An important characteristic of Twitter is its real-time nature. For example, when an earthquake occurs, people make many Twitter posts (tweets) related to the earthquake, which enables detection of earthquake occurrence promptly, simply by observing the tweets. As described in this paper, we investigate the real-time interaction of events such as earthquakes, in Twitter, and propose an algorithm to monitor tweets and to detect a target event. To detect a target event, we devise a classifier of tweets based on features such as the keywords in a tweet, the number of words, and their context. Subsequently, we produce a probabilistic spatiotemporal model for the target event that can find the center and the trajectory of the event location. We consider each Twitter user as a sensor and apply Kalman filtering and particle filtering, which are widely used for location estimation in ubiquitous/pervasive computing. The particle filter works better than other compared methods in estimating the centers of earthquakes and the trajectories of typhoons. As an application, we construct an earthquake reporting system in Japan. Because of the numerous earthquakes and the large number of Twitter users throughout the country, we can detect an earthquake by monitoring tweets with high probability (96% of earthquakes of Japan Meteorological Agency (JMA) seismic intensity scale 3 or more is detected). Our system detects earthquakes promptly and sends e-mails to registered users. Notification is delivered much faster than the announcements that are broadcast by the JMA.

E. Text Detection and Recognition on Traffic Panels from Street-Level Imagery Using Visual Appearance

Traffic sign detection and recognition has been thoroughly studied for a long time. However, traffic panel detection and recognition still remains a challenge in computer vision due to its different types and the huge variability of the information depicted in them. This paper presents a method to detect traffic panels in street level images and to recognize the information contained on them, as an application to intelligent transportation systems (ITS). The main purpose can be to make an automatic inventory of the traffic panels located in a road to support road maintenance and to assist drivers. Our proposal extracts local descriptors at some interest key points after applying blue and white color segmentation. Then, images are represented as a “bag of visual words” and classified using Naïve Bayes or support vector machines. This visual appearance categorization method is a new approach for traffic panel detection in the state of the art. Finally, our own text detection and recognition method is applied on those images where a traffic panel has been detected, in order to automatically read and save the information depicted in the panels. We propose a language model partly based on a dynamic dictionary for a limited geographical area using a reverse geo coding service. Experimental results on real images from Google Street View prove the efficiency of the proposed method and give way to using street-level images for different applications on ITS.

IV. REAL-TIME DETECTION OF TRAFFIC EVENTS The developed system was installed and tested for the real-time monitoring of several areas of the Italian road network, by means of the analysis of the Twitter stream coming from those areas. The aim is to perform a continuous monitoring of frequently busy roads and highways in order to detect possible traffic events in real-time or even in advance with respect to the traditional news media. The system is implemented as a service of a wider service-oriented platform to be developed in the context of the SMARTY project. The service can be called by each user of the platform, who desires to know the traffic conditions in a certain area. In this section, we aim to show the effectiveness of our system in determining traffic events in short time.

TABLE I: Classification Results On The 2-Class Dataset (Best Values In Bold)

Classifier	Accuracy (%)	Precision (%) by class		Recall (%) by class		F_1-score (%) by class	
		Traffic	*Non-traffic*	*Traffic*	*Non-traffic*	*Traffic*	*Non-traffic*
SVM	**95.75**	**95.3**	**96.3**	**96.5**	**95.0**	**95.8**	**95.7**
C4.5	95.15	94.4	96.1	96.1	94.2	95.2	95.1
1NN	91.87	93.2	91.2	90.9	93.3	92	92.2
3NN	91.69	93.3	90.3	89.9	93.5	91.5	91.8
5NN	91.61	93.7	89.9	89.3	93.9	91.4	91.8
NB	90.56	93.2	88.4	87.7	93.4	90.3	90.8
PART	94.66	94.1	95.4	95.3	94.0	94.7	94.6

TABLE II: Results of the Classification of Tweets in Other Works in the Literature

Classification algorithm	Considered tweets' classes	Classification results		Dataset	
		Measure	Value (%)	Size	Class balancing
NLP analysis [12]	*traffic-related* vs. *non-traffic-related*	Accuracy	91.75	1249 tweets	no
		Precision	91.39		
		Recall	87.53		
NLP analysis [31]	*CDE-related* vs. *non-CDE-related*	Accuracy	80.00	-	-
SVM [24]	*incident-related* vs. *non-incident-related*	Accuracy	89.06	640 tweets	yes
		Precision	89.10		
		Recall	89.10		
		F-score	89.10		
NB [24]	*incident-related* vs. *non-incident-related*	Accuracy	86.25	640 tweets	yes
		Precision	87.30		
		Recall	86.30		
		F-score	86.20		
RIPPER [24]	*incident-related* vs. *non-incident-related*	Accuracy	85.93	640 tweets	yes
		Precision	86.80		
		Recall	85.90		
		F-score	85.90		
SVM [9]	*heavy-traffic* vs. *non-heavy-traffic*	Precision	87.00	120 tweets	-
		Recall	67.00		
		F-score	75.00		

International Journal of Research
Available at https://edupediapublications.org/journals

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 10
September 2017

TABLE III: Classification Results on The 3-Class Dataset (Best Values In Bold)

Classifier	Accuracy (%)	Precision (%) by class			Recall (%) by class			F_1-score (%) by class		
		Traffic cong. or crash	*Traffic due to ext. event*	*Non-traffic*	*Traffic cong. or crash*	*Traffic due to ext. event*	*Non-traffic*	*Traffic cong. or crash*	*Traffic due to ext. event*	*Non-traffic*
SVM	**88.89**	**81.4**	**93.5**	**93.9**	**92.8**	85.9	88.0	**86.6**	**89.5**	**90.8**
C4.5	86.03	77.6	90.4	92.5	89.0	80.3	**88.9**	82.8	84.9	90.5
1NN	80.53	73.3	84.6	86.0	79.3	77.9	84.5	76.2	80.7	85.0
3NN	81.13	70.3	93.8	85.5	85.7	71.6	86.2	77.2	80.7	85.6
5NN	80.28	68.2	94.6	87.5	90.1	65.0	85.7	77.5	76.8	86.4
NB	81.23	75.9	77.0	94.0	75.3	**86.1**	81.1	75.4	81.4	86.9
PART	85.04	77.1	89.5	91.5	87.8	79.9	87.4	81.8	84.1	89.4

TABLE IV: Results of the Wilcoxon Signed-Rank Test on The Accuracies Obtained On The Test Set For The 2-Class Dataset

Comparison	R^+	R^-	p-value	Hypothesis ($\alpha = 0.05$)
SVM vs. C4.5	170	40	$1.4 \cdot 10^{-2}$	Rejected
SVM vs. 1NN	210	0	$1.9074 \cdot 10^{-6}$	Rejected
SVM vs. 3NN	210	0	$1.9074 \cdot 10^{-6}$	Rejected
SVM vs. 5NN	210	0	$1.9074 \cdot 10^{-6}$	Rejected
SVM vs. NB	210	0	$1.9074 \cdot 10^{-6}$	Rejected
SVM vs. PART	167	23	$2 \cdot 10^{-3}$	Rejected

TABLE V: Results of the Wilcoxon Signed-Rank Test on the Accuracies Obtained On The Test Set For The 3-Class Dataset

Comparison	R^+	R^-	p-value	Hypothesis ($\alpha = 0.05$)
SVM vs. C4.5	197	13	$1.6784 \cdot 10^{-4}$	Rejected
SVM vs. 1NN	210	0	$1.9074 \cdot 10^{-6}$	Rejected
SVM vs. 3NN	210	0	$1.9074 \cdot 10^{-6}$	Rejected
SVM vs. 5NN	210	0	$1.9074 \cdot 10^{-6}$	Rejected
SVM vs. NB	208	2	$5.722 \cdot 10^{-6}$	Rejected
SVM vs. PART	202.5	7.5	$4.196 \cdot 10^{-5}$	Rejected

We just present some results for he2-class problem. For the setup of the system, we have employed as training set the overall dataset. We adopt only the best performing classifier, i.e., the SVM classifier. During the learning stage, we identified Q = 3227 features, which were reduced to F = 582 features after the feature selection step. The system continuously performs the following operations: it i) fetches, with a time frequency of z minutes, tweets originated from a given area, containing the keywords resulting from CondA, ii) performs a real-time classification of the fetched tweets, iii) detects a possible traffic-related event, by analyzing the traffic class tweets from the considered area, and, if needed, sends one or more traffic warning signals with increasing intensity for that area. More in detail, a first low-intensity warning signal is sent when m traffic class tweets are found in the considered area in the same or in subsequent temporal windows. Then, as the number of traffic class tweets grows, the warning signal becomes more reliable, thus more intense. The value of m was set based on heuristic considerations, depending, e.g., on the traffic density of the monitored area. In the experiments we set m = 1. As regards the fetching frequency z, we heuristically found that z = 10 minutes represents a good compromise between fast event detection and system scalability. In fact, z should be set depending on the number of monitored areas and on the volume of tweets fetched. With the aim of evaluating the effectiveness of our system, we need that each detected traffic-related event is appropriately validated.

Validation can be performed in different ways which include: i) direct communication by a person, who was present at the moment of the event, ii) reports drawn up by the police and/or local administrations (available only in case of incidents), iii) radio traffic news; iv) official real-time traffic news web sites; v) local newspapers (often the day after the event and only when the event is very significant). Direct communication is possible only if a person is present at the event and can communicate this event to us. Although we have tried to sensitize a number of users, we did not obtain an adequate feedback. Official reports are confidential: police and local administrations barely allow accessing to these reports, and, when this permission is granted, reports can be consulted only after several days. Radio traffic news is in general quite precise in communicating traffic-related events in real time. Unfortunately, to monitor and store the events, we should dedicate a person or adopt some tool for audio analysis. We realized however that the traffic-related events communicated on the radio are always mentioned also in the official real-time traffic news web sites. Actually, on the radio, the speaker typically reads the news reported on the web sites. Local newspapers focus on local traffic-related events and often provide events which are not published on official traffic news web sites. Concluding, official real-time traffic

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 10
September 2017

news web sites and local newspapers are the most reliable and effective sources of information for traffic-related events. Thus, we decided to analyze two of the most popular real-time traffic news web sites for the Italian road network, namely "CCISS Viaggiare informati",8 managed by the Italian government Ministry for infrastructures and transports, and "Autos trade per l'Italia",9 the official web site of Italian highway road network. Further, we examined local newspapers published in the zones where our system was able to detect traffic-related events.

Actually, it was really difficult to find realistic data to test the proposed system, basically for two reasons: on the one hand, we have realized that real traffic events are not always notified in official news channels; on the other hand, situations of traffic slowdown may be detected by traditional traffic sensors but, at the same time, may not give rise to tweets. In particular, in relation to this latter reason, it is well known that drivers usually share a tweet about a traffic event only when the event is unexpected and really serious, i.e., it forces to stop the car. So, for instance, they do not share a tweet in case of road works, minor traffic difficulties, or usual traffic jams (same place and same time). In fact, in correspondence to minor traffic jams we rarely find tweets coming from the affected area. We have tried to build a meaningful set of traffic events, related to some major Italian cities, of which we have found an official confirmation. The selected set includes events correctly identified by the proposed system and confirmed via official traffic news web sites or local newspapers. The set of traffic events, whose information is summarized in Table VI, consists of 70 events detected by our system. The events are related both to highways and to urban roads, and were detected during September and early October 2014.

Table VI shows the information about the event, the time of detection from Twitter's stream fetched by our system, the time of detection from official news websites or local newspapers, and the difference between these two times. In the table, positive differences indicate a late detection with respect to official news web sites, while negative differences indicate an early detection. The symbol "-" indicates that we found the official confirmation of the event by reading local newspapers several hours late. More precisely, the system detects in advance 20 events out of 59 confirmed by news web sites, and 11 events confirmed the day after by local newspapers. Regarding the 39 events not detected in advance we can observe that 25 of such events are detected within 15 minutes from their official notification, while the detection of the remaining 14 events occurs beyond 15 minutes but within 50 minutes. We wish to point out, however, that, even in the cases of late detection, our system directly and explicitly notifies the event occurrence to the drivers or passengers registered to the SMARTY platform, on which our system runs. On the contrary, in order to get traffic information, the drivers or passengers usually need to search and access the official news websites, which may take some time and effort, or to wait for getting the information from the radio traffic news.

As future work, we are planning to integrate our system with an application for analyzing the official traffic news web sites, so as to capture traffic condition notifications in real-time. Thus, our system will be able to signal traffic-related events in the worst case at the same time of the notifications on the web sites. Further, we are investigating the integration of our system into a more complex traffic detection infrastructure. This infrastructure may include both advanced physical sensors and social sensors such as streams of tweets. In particular, social sensors may provide a low-cost wide coverage of the road network, especially in those areas (e.g., urban and suburban) where traditional traffic sensors are missing.

V. CONCLUSION AND FUTURE SCOPE

- In this paper, we have proposed a system for real-time detection of traffic-related events from Twitter stream analysis.
- The system, built on a SOA, is able to fetch and classify streams of tweets and to notify the users of the presence of traffic events.

- Furthermore, the system is also able to discriminate if a traffic event is due to an external cause, such as football match, procession and manifestation, or not.

Future Scope: As future work, we are planning to integrate our system with an application for analyzing the official traffic news web sites, so as to capture traffic condition notifications in real-time. Thus, our system will be able to signal traffic-related events in the worst case at the same time of the notifications on the web sites. Further, we are investigating the integration of our system into a more complex traffic detection infrastructure. This infrastructure may include both advanced physical sensors and social sensors such as streams of tweets. In particular, social sensors may provide a low-cost wide coverage of the road network, especially in those areas (e.g., urban and suburban) where traditional traffic sensors are missing.

VI. REFERENCES

[1] Eleonora D'Andrea, Pietro Ducange, Beatrice Lazzerini, Member, IEEE, and Francesco Marcelloni, Member, IEEE, "Real-Time Detection of Traffic From Twitter Stream Analysis", IEEE Transactions on Intelligent Transportation Systems, Vol. 16, No. 4, August 2015.

[2] F. Atefeh and W. Khreich, "A survey of techniques for event detection in Twitter," Comput. Intell., vol. 31, no. 1, pp. 132–164, 2015.

[3] P. Ruchi and K. Kamalakar, "ET: Events from tweets," in Proc. 22nd Int. Conf. World Wide Web Comput., Rio de Janeiro, Brazil, 2013, pp. 613–620.

[4] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in Proc. 7th ACM SIGCOMM Conf. Internet Meas., San Diego, CA, USA, 2007, pp. 29–42.

[5] G. Anastasi et al., "Urban and social sensing for sustainable mobility in smart cities," in Proc. IFIP/IEEE Int. Conf. Sustainable Internet ICT Sustainability, Palermo, Italy, 2013, pp. 1–4.

[6] A. Rosi et al., "Social sensors and pervasive services: Approaches and perspectives," in Proc. IEEE Int. Conf. PERCOM Workshops, Seattle, WA, USA, 2011, pp. 525–530.

[7] T. Sakaki, M. Okazaki, and Y.Matsuo, "Tweet analysis for real-time event detection and earthquake reporting system development," IEEE Trans. Knowl. Data Eng., vol. 25, no. 4, pp. 919–931, Apr. 2013.

[8] J. Allan, Topic Detection and Tracking: Event-Based Information Organization. Norwell, MA, USA: Kluwer, 2002.

[9] K. Perera and D. Dias, "An intelligent driver guidance tool using location based services," in Proc. IEEE ICSDM, Fuzhou, China, 2011, pp. 246–251.

[10] T. Sakaki, Y. Matsuo, T. Yanagihara, N. P. Chandrasiri, and K. Nawa, "Real-time event extraction for driving information from social sensors," in Proc. IEEE Int. Conf. CYBER, Bangkok, Thailand, 2012, pp. 221–226.

[11] 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.

[12] A. Gonzalez, L. M. Bergasa, and J. J. Yebes, "Text detection and recognition on traffic panels from street-level imagery using visual appearance," IEEE Trans. Intell. Transp. Syst., vol. 15, no. 1, pp. 228–238, Feb. 2014.