Computers, Materials & Continua
DOI:10.32604/cmc.2020.012057
Article

Tech Science Press

Autonomous Eyewitness Identification by Employing Linguistic Rules for Disaster Events

Sajjad Haider* **and Muhammad Tanvir Afzal**

Capital University of Science and Technology, Islamabad, 44000, Pakistan
*Corresponding Author: Sajjad Haider. Email: sajjadhaidermalik@gmail.com
Received: 12 June 2020; Accepted: 03 August 2020

Abstract: Social networking platforms provide a vital source for disseminating information across the globe, particularly in case of disaster. These platforms are great mean to find out the real account of the disaster. Twitter is an example of such platform, which has been extensively utilized by scientific community due to its unidirectional model. It is considered a challenging task to identify eyewitness tweets about the incident from the millions of tweets shared by twitter users. Research community has proposed diverse sets of techniques to identify eyewitness account. A recent state-of-the-art approach has proposed a comprehensive set of features to identify eyewitness account. However, this approach suffers some limitation. Firstly, automatically extracting the feature-words remains a perplexing task against each feature identified by the approach. Secondly, all identified features were not incorporated in the implementation. This paper has utilized the language structure, linguistics, and word relation to achieve automatic extraction of feature-words by creating grammar rules. Additionally, all identified features were implemented which were left out by the state-of-the-art model. A generic approach is taken to cover different types of disaster such as earthquakes, floods, hurricanes, and wildfires. The proposed approach was then evaluated for all disaster-types, including earthquakes, floods, hurricanes, and fire. Based on the static dictionary, the Zahra et al. approach was able to produce an F-Score value of 0.92 for Eyewitness identification in the earthquake category. The proposed approach secured F-Score values of 0.81 in the same category. This score can be considered as a significant score without using a static dictionary.

Keywords: Grammar rules; social media; eyewitness identification; disaster response

1 Introduction

In today's digital age social media platforms like Facebook, Twitter, and Instagram are widely used for day to day activities. People across the globe harness these platforms to share information, ideas, opinions, and reviews regarding various items or topics [1,2]. Twitter is most widely used among these social media platforms, due to its unique unidirectional relationship model. Twitter handles 500+ million tweets every day,

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

which are posted or re-tweeted by 134 million active users per day. This volume is believed to be increased by 30% a year [3].

Twitter's word wide popularity has attracted the research community to discover implicit and explicit information from it. There has been study on Twitter for targeted news recommendations, disaster and emergency alerts, response systems, and advertising, etc. [1]. Twitter is also deemed as a potential medium of *breaking news*. It has proved itself as a news breaker, 85% of trending topics on twitter are news headlines [4].

The question arises, how Twitter does that? The answer is through the eyewitness' tweets posted by the user. Some of the events wherein Twitter has proved its capacity of breaking the news are outlined below:

- A news agency found a tweet posted by a passenger of Delta Aircraft flight which made an emergency landing due to the potential engine failure on a remote island of Alaska [5].
- An eyewitness from Hudson Bay tweeted about the New York Airplane crash, and Daily Telegraph broke the news as a headline [6].
- There were half a dozen tweets available on Twitter a minute before USGS [7] recorded the time about the California Earthquake [8].
- The eyewitness tweets were available about the bombing incident of Boston well before the coverage of the news channels [9].
- The attack on Westgate Shopping Mall in Nairobi, Kenya was on Twitter, thirty-three minutes before being reported on the TV news Channels [10].

Literature reviews suggest identification of eyewitnesses' is a vital task as the information shared by an eyewitness can better explain the state and severity of the event [11]. Furthermore, the emergency services or agencies responding to any disastrous event tend to rely on the information provided by eyewitness account. Emergency services have to make quick and effective measures to control the situation therefore credible information is vital. The scientific community has been focusing on the eyewitness tweets identification for several years [11,12]. The state-of-the-art approaches have adopted diversified techniques and feature sets to identify eyewitness tweet from a large pool of text.

A recent study by Zahra et al. [11] identified a comprehensive list of features to identify eyewitness tweets. These features were applied to a set of 2000 tweets for training and 8000 tweets for testing. The authors have asked domain-experts to manually identify thirteen features for eyewitness identification as described below:

- *Feature-1: Reporting small details of surroundings:* e.g., window shaking, water in basement
- *Feature-2: Words indicating perceptual senses:* e.g., seeing, hearing, feeling
- *Feature-3: Reporting impact of disaster:* e.g., raining, school canceled, flight delayed
- *Feature-4: Words indicating intensity of disaster:* e.g., intense, strong, dangerous, big
- *Feature-5: First person pronouns and adjectives:* e.g., i, we, me
- *Feature-6: Personalized location markers:* e.g., my office, our area
- *Feature-7: Exclamation and question marks:* e.g., !, ?
- *Feature-8: Expletives:* e.g., wtf, omg, s**t
- *Feature-9: Mention of a routine activity:* e.g., sleeping, watching a movie
- *Feature-10: Time indicating words:* e.g., now, at the moment, just
- *Feature-11: Short tweet length:* e.g., one or two words
- *Feature-12: Caution and advice for others:* e.g., watch out, be careful
- *Feature-13: Mention of disaster locations:* e.g., area and street name, directions

A tweet posted by an eyewitness user is shown in Fig. 1. The words are tagged with appropriate features identified by Zahra et al. [11]. The identified features are marked in different colors with their corresponding labels in Fig. 1.

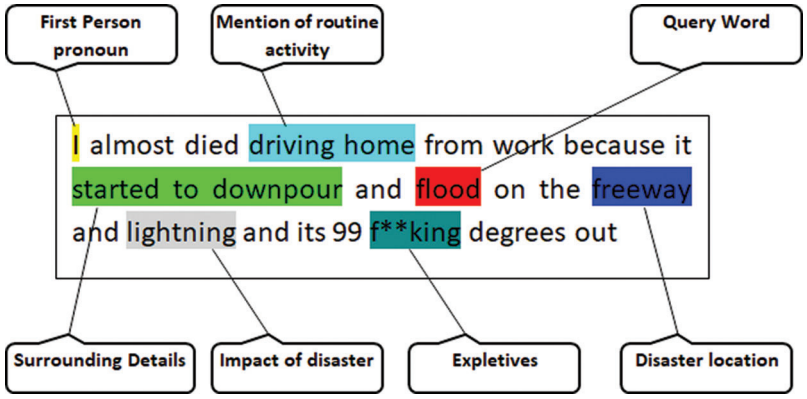

Figure 1: Example tweet explained with domain-expert features

Zahra et al. [11] adopted the manually created dictionary to identify the eyewitness. The authors have dropped the implementation part of some identified features like "Mention of disaster locations", "Reporting small details of surroundings", and "Personalized location markers" due to their implementation complexity.

In this study, we define a set of grammar rules which recognizes the feature-words for all the identified features proposed by Zahra et al. [11]. Similarly, we have also defined the grammar rules for those features which were dropped due to implementation complexity. Additionally our proposed approach is fully automated wherein no human interaction is required for feature-words extraction.

The proposed approach has exploited the language structure, linguistic features, and existing relationship among words of a sentence in order to explain the context. Grammar rules are defined to automatically extract the feature-words from the tweet content. Furthermore, the proposed approach has been evaluated for disasters like earthquakes, floods, hurricanes, and wildfires.

We have evaluated the proposed approach on a benchmark dataset. The results were compared with Zahra et al. [11] approach in terms of precision, recall, and f-measure. The proposed approach achieved the f-score of 0.81 for the earthquake events, which was comparable to the manual approach of Zahra et al. [11] f-score of 0.92.

In the next section, we shall explore some of the exciting and innovative methods for the identification of eyewitnesses. The structure of our proposed approach and a comprehensive discussion of all the features are illustrated in Section 3. Evaluation results and their comparison with Zahra et al. [11] approach are explained in Section 4. The conclusion of our approach is presented in Section 5.

2 Literature Review

Twitter is one of the most commonly used social media wherein people share their opinions and experiences about different news and events happening around them. In the event of a natural disaster, people search for the latest news and real-time content on Twitter, please see Kryvasheyeu et al. [13–16].

The extraction of useful information from the users' tweets is a critical task, especially when it comes to the disaster events. The disaster management system requires precise and accurate information to control the situation in a prompt manner. Twitter and disaster management mainly focus on the user-provided

information for disaster response and relief [17,18]. The Haiti earthquake was utilized by Meier [19], and Ostermann et al. [20] exploited the forest fires events.

The accounts of twitter users and the relational network are used by Truelove et al. [21] to tackle the eyewitness accounts. A predefined list of query words is tested for their presence while pre-processing the content to identify the target tweet. However, this study has not presented any characteristics to identify eyewitness tweets [21]. In comparison to Truelove et al. [21,22] propose a list of five linguistic features to identify the eyewitness tweets. Fang et al. [23] presented a hybrid approach by adopting linguistic features with meta-features, like the application used for reporting, for the identification of eyewitness reports. Fang et al. propose five stylistic features with the set of five linguistic features for the identification of eyewitness tweets. Tanev et al. [24] have also proposed a hybrid approach which uses three stylistic features and three linguistic features with Twitter meta-data for eyewitness identification.

Various applications are developed to crawl the real-time tweets to facilitate disaster relief organizations, like Tweet Tracker [25], Twitcident [26], AIDR [27], and Scatter Blogs for situational awareness [28]. The credibility of crowd sourced based generated data suffers the quality assurance for objectivity and truthfulness of information [29] as required by the real-life disaster management systems.

The organizations which need to respond to a disaster always search for a credible news source or eyewitness tweet [30]. The person witnessing the event can truthfully provide details like intensity, effect, casualties, and other information regarding the incident. Diakopoulos et al. [31] presented a study of eyewitness report identification out of millions of tweets related to journalism. A similar study on natural disasters and criminal justice was presented by Olteanu et al. [32] in 2015. Location information of the users is utilized to access remote and local users for a disaster event by Kumar et al. [33] in 2013. Morstatter et al. [34] proposed a technique to identify set of features for automatic classification of disaster events, based on language and the linguistic patterns found in the tweet content. Kumar et al. [33,34] studies have used the location and language information to capture the source whereas an eyewitness account is not taken into consideration.

Truelove et al. [21] proposed a conceptual model to identify the witness account and its related impact and relayed accounts and evaluated for various events. In 2016 Doggett et al. [22] proposed the linguistic feature sets for eyewitness and non-eyewitness categories to identify eyewitness tweets in the event of a disaster. The authors have identified five features of eyewitness and four for non-eyewitness categories. Fang et al. [23] developed an eyewitness identification task by defining a set of linguistics and meta-features. The authors have used the word dictionary of LIWC [35] and OpenCalais API [36] for topic classification. Tanev et al. [24] in 2017 defined the eyewitness features using the lexical dimension, stylistic dimension, metadata, and Semantics. None of the above techniques developed their classifiers through a combination of expert-driven engineering as the expert-driven features related to disaster events can be helpful in the identification of true eyewitness sources.

A recent approach by Zahra et al. [11] in 2019 proposed a technique to categories features of eyewitness types from eyewitness reports using expert-driven engineering of frequent natural disasters. The authors identified the feature set associated with eyewitness and called them the domain-expert features. The authors experimented with the textual features (bag-of-words) combined with the domain-expert features for the classification of eyewitness tweets.

We have critically analyzed all of the above studies which led us to identify the important problem discussed in this paper. The summary of eyewitness features identification techniques is illustrated in Tab. 1. The table demonstrates the methodology, results and shortcomings of [11] in tweet identification.

We have formed a complete list of features which have been used by different studies to tackle the intended tweets. Tab. 2 below shows the list of features against the approaches which have incorporated.

Table 1: Summary of eyewitness features identification techniques

Literature	Idea	Methodology	Achieved	Weaknesses
Truelove et al. [21]	This research exploited the idea of identifying the witness account and related impact accounts.	They proposed a model to identify the witness account, related, and relayed impact accounts tweets of Bush-fire of 18-19th Feb 2013, Hume Freeway road connecting Melbourne and Sydney.	Achieved 77% results of observations of smoke, but the effect of the event is not a true eyewitness due to the distance.	Require Location and Network info, metadata, manual pre-processing. No eyewitness features identify.
Doggett et al. [22]	Propose a filter based technique for identifying posts from eyewitnesses using predefined keywords.	The author identified five eyewitness linguistic features and used them as filters to identify eyewitness events.	Eyewitness filters Avg. accuracy 62%.	Require Geo-location info of the Twitter user. Identifies eyewitness events, not the eyewitness.
Fang et al. [23]	Propose a hybrid approach for automatic identification of eyewitness during emergency events.	Eyewitness identification task by defining a set of linguistics and meta-features. Used OpenCalais API to label event and LIWC dictionary to identify events related keywords.	Average F-score of 89.7%.	Used a manually created list of terms. The model requires location and language information for identification.
Tanev et al. [24]	It presents grammar learning and natural language algorithms for event detection.	To determine the disaster effect, unsupervised algorithms, and evaluated for event detection from news articles.	Achieved 42% Precision and Recall of 66%	Domain-specific news data for event detection.
Zahra et al. [11]	Manually investigates sources on tweets related to eyewitnesses and classifies them into thirteen features for identification.	Domain experts manually analyzed a set of tweets for feature identification from the content. Identified feature words for each feature type.	F-Score of 0.917	Manual implementation. Failed to Implement all features.

Table 2: Summary of identified eyewitness features

Identified eyewitness feature	Doggett et al. [22]	Fang et al. [23]	Tanev et al. [24]	Zahra et al. [11]
Reporting small details of surroundings	–	Y	–	Y
Words indicating perceptual senses	–	Y	–	Y
Reporting impact of disaster	–	–	Y	Y
Words indicating the intensity of the disaster	–	–	–	Y
First-person pronouns and adjectives	Y	Y	Y	Y
Personalized location markers	Y	–	–	Y
Exclamation and question marks	Y	–	–	Y
Expletives	Y	–	–	Y
Mention of a routine activity	–	–	–	Y
Time indicating words	Y	Y	–	Y
Short tweet-length	–	Y	Y	Y
Caution and advice for others	–	–	–	Y
Mention of disaster locations	–	–	–	Y
All Caps	–	–	Y	–
Contains media	–	Y	–	–
URL	–	Y	Y	–
Journalist account involved	–	Y	–	–
Type of client application used	–	Y	Y	–
Mentions or Hash-tags	–	Y	–	–
Word Embedding (Semantic word relationship)	–	Y	–	–

We have identified two main issues which have been overlooked by Zahra et al. approach on tweets identification. Firstly, most part of the study has used a static dictionary, which is not scale-able for different domains and unseen tweets. Secondly, the implementation of the approach is not practical when millions of tweets need to be processed in real-time. The manual extraction of features with the help of a domain expert using a predefined list is time consuming and updating new data set in real-time is cumbersome.

In the light of above limitation, we have proposed an automatic tweet identification system. It can intelligently processes millions of tweets in real-time without requiring domain experts to build static dictionaries for different domains.

3 Methodology

A comprehensive discussion on the proposed methodology is illustrated in this section. The overall approach is described in Fig. 2. Rule identification for feature extraction module is described in Fig. 3.

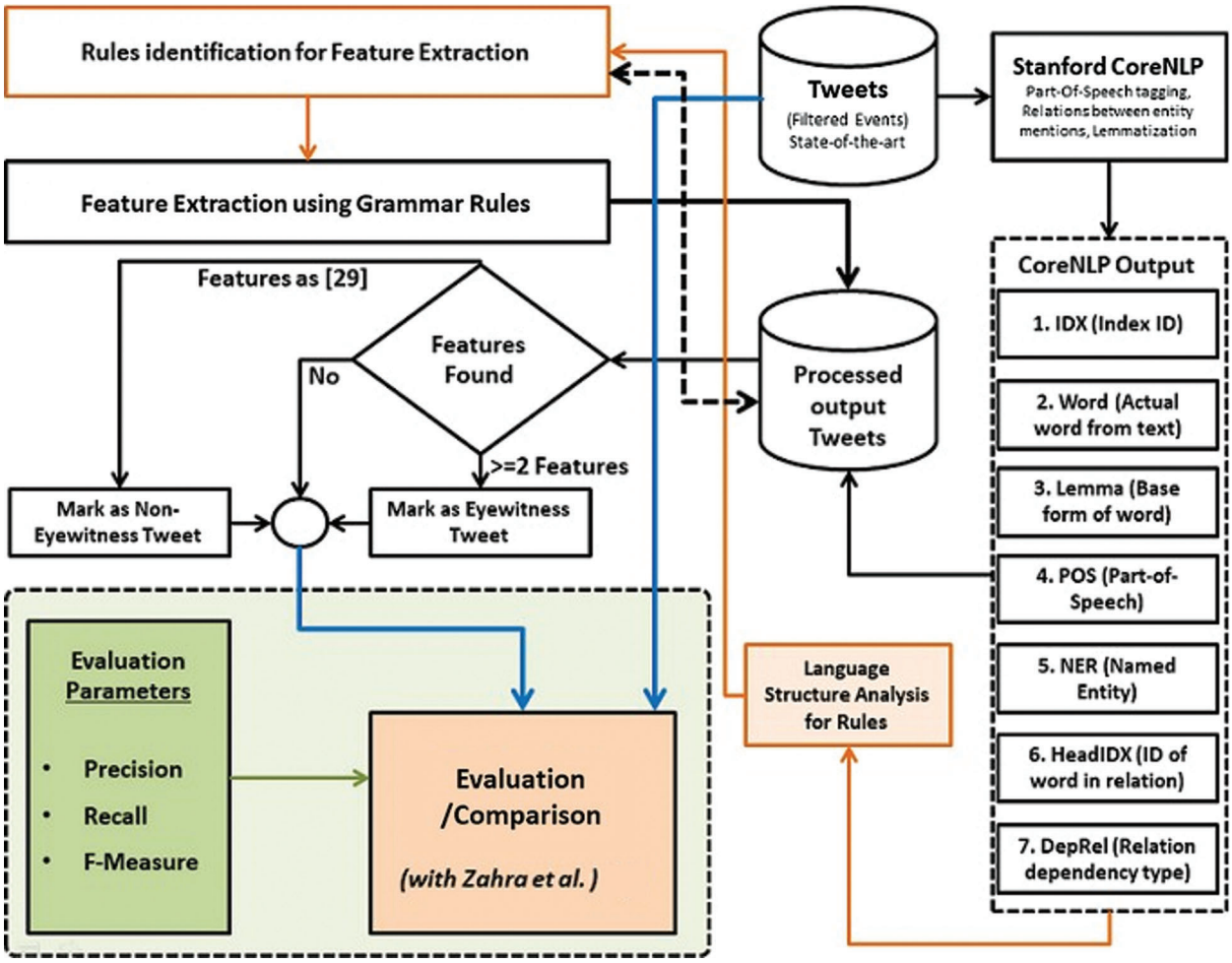

Figure 2: Proposed methodology

The data employed to implement the proposed methodology contains tweets containing words such as: "earthquake", "flood", "hurricane", "wildfire". This data set has also been employed and publicly released by Zahra et al. [11] approach. We have acquired the data from CrisisNLP repository.

Firstly, the tweets are pre-processed to remove the noise. Thereafter, they are parsed and annotated Part-of-Speech (POS) tagger. The tagger also provides lemmatization and entity-relationship details. We have used the CoreNLP [37] tool, which labels each word to its respective POS and also identifies the relationship between the words of each sentence. The output of CoreNLP is then evaluated manually by the language structure to identify the grammar rules for each identified feature as shown in Fig. 3.

The grammatical rules suggested for each function are then added to the processing of data, and the findings are compiled. Now, the next step is to identify tweets as "eyewitness", "non-eyewitness", or "Unknown" source. For this, we have searched for those tweets that contain at least two features of the specified set. If the tweet includes two features then it is categorized as "eyewitness", if it does not have any explicit feature but includes non-eyewitness features from Doggett et al. [22], it is classified as "non-eyewitness". Tweets containing no identified feature are categorized as "unknown". The following subsections provide a brief explanation of each process involved in the proposed methodology from data set selection to results evaluation.

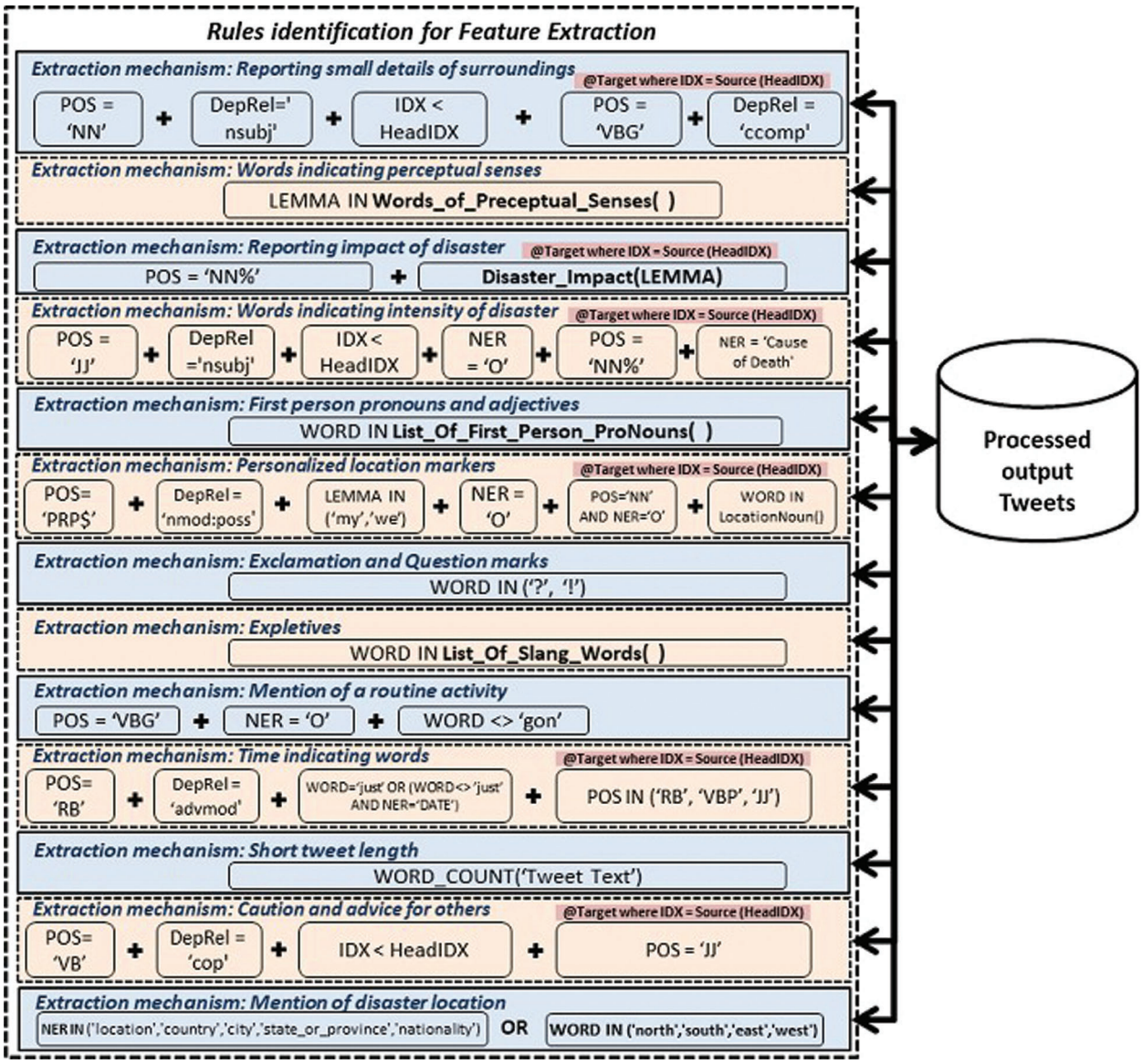

Figure 3: Proposed methodology (rule identification for feature extraction)

3.1 Dataset

The dataset employed by Zahra et al. [38] contains tweets from July-2016 to May-2018 utilizing the Twitter Streaming API. This data contains 2000 tweets, which were tweeted from 1st to 28th August 2017 with focused keywords including earthquake, foreshock, aftershock, hurricane, fire, forest fire, flood, etc. These tweets were used for manual analysis by Zahra et al. [11] approach. Another data set, which is also employed in our proposed study, comprises 8000 tweets picked from random collection based on same keywords and annotation. This data is collected by using Twitter Streaming API [11]. It contains annotations given by the authors.

3.2 Pre-Processing

Tweets are generally written in natural languages where grammar and other writing rules are not followed. To remove redundant information from the tweets, we have pre-processed this data. Tweets pre-

processing involves the removal of HTML tags, hash-tags, extra white spaces, special symbols, etc. After the pre-processing procedure tweets are parsed.

3.3 Parsing of Tweets (Stanford CoreNLP)

The Part-Of-Speech (POS) tagging or grammatical tagging is the process wherein each word in a text-*corpus* is tagged with its corresponding part of speech, based on both the definition and the context. The Stanford NER (Named Entity Recognition) has been proven useful in attaining high F-measure scores in contemporary approaches [11,38]. Therefore, we have also used the Stanford CoreNLP tool. The annotations of the tweets are also stored in the same database which contains all the tweets to be processed in order to carryout comparison.

For instance, consider the tweet text; "I felt something, and my door was shaking", and its output is obtained by using the Stanford CoreNLP, as shown in Tab. 3.

Table 3: Stanford CoreNLP output

IDX	Word	Lemma	POS	NER	HeadIDX	DepRel
1	I	I	PRP	O	2	nsubj
2	Felt	Feel	VBD	O	0	ROOT
3	Something	something	NN	O	8	nsubj
4	And	And	CC	O	3	cc
5	My	My	PRP$	O	6	nmod:poss
6	Door	Door	NN	O	3	conj
7	Was	Be	VBD	O	8	aux
8	shaking	Shake	VBG	O	2	ccomp

In Tab. 3, the first column "IDX" is the index ID of a word in a sentence of the tweet. The second column is the actual Word field. The third column field is titled as "Lemma" as it has a base word of the focused term generated by the lemmatization process. The forth column "POS" contains the corresponding Part-of-speech of the focused term. The fifth column titled "NER" contains the Named-Entity tag. Whereas columns sixth and seventh both contain the related information of the focused term. The sixth column of HeadIDX contains the Index-ID of the related word or zero (identifying the relationship between words). The last column contains the Dependency relation to the second word. Stanford provided the complete typed dependencies manual in 2008 and revised it in 2016 [39].

3.4 Feature Extraction (Grammar Rule-Based)

In-depth analysis of the annotated dataset revealed that language structure and the relationship between different words of a sentence could be exploited for automatic extraction of the features from tweet content. Unlike manually created dictionaries or lists used by the existing Zahra et al. [11] approach, we have proposed grammar rules for almost every feature. Moreover, our approach also enjoys created rules for features that were dropped by the author due to their implementation complexity.

We proposed grammar-based rules for maximum features, but few features like "Words indicating perceptual senses", "First-person pronouns and adjectives", "Exclamation and question marks", and "Expletives" are somehow limited to be implemented only by using a predefined list of words. Similarly, for "Expletives" words, no rules can be defined due to the scope of this feature and again have to limit

our implementation to dictionaries rather than defining grammar rules. Out of thirteen features, we have to use dictionaries for four features and a word count approach for one identified feature. The following subsections briefly discuss each feature and its proposed Grammar Rules or the limitations of using lists.

3.4.1 Feature-1: Reporting Small Details of Surroundings

The first feature identified by Zahra et al. [11] approach is the identification of "small surrounding details" reported by tweet authors. The feature is not implemented by Zahra et al. [11] approach, and states that it "proved too abstract to be implemented". From a human perspective, this is an important feature to understand the context. The importance of the feature is also exploited by Fang et al. [23]. We have created grammar rules for the extraction of this feature. Working on the grammar rule is explained in Tab. 4, the data is the same as we have described in Tab. 3.

Table 4: Stanford CoreNLP output (working)

IDX	Word	Lemma	POS	NER	HeadIDX	DepRel
1	I	I	PRP	O	2	nsubj
2	Felt	Feel	VBD	O	0	ROOT
3	Something	something	NN	O	8	nsubj
4	And	And	CC	O	3	cc
5	My	My	PRP$	O	6	nmod:poss
6	Door	Door	NN	O	3	conj
7	Was	Be	VBD	O	8	aux
8	Shaking	Shake	VBG	O	2	ccomp

The relationship among tagged words is shown in Tab. 4. In this text, "something shaking" is the feature that explains the surrounding detail. By using a bag-of-words technique, it is not possible to extract this feature word.

We have achieved successful extraction of such features, using grammar rules. It is illustrated in Tab. 4 where we can see that at token 3 (IDX = 3), we have a noun (NN) and it has a dependency relation (DepRel = "nsubj") with token 8 (IDX = 8) that we can link from HeadIDX value (HeadIDX = 8). The target token is a verb (POS = "VBG") and has the dependency relation (DepRel = "ccomp"). By implementing this language rule, we can identify the feature "something shaking", that explains the surrounding details. For implementation purpose, we have proposed a rule as follows:

RULE-1; POS= 'NN' and DepRel= 'nsubj' and IDX <HeadIDX and NER<>

'CAUSE_OF_DEATH' and OnTarget(POS= 'VBG' and DepRel= 'ccomp')

3.4.2 Feature-2: Words Indicating Perceptual Senses

The words that come under this category related to a limit list like "hearing, seeing". For its implementation, we have used the technique of "Lemma matching" of each token with a list of lemma words of "hearing, seeing". A list of such words is available online [35]; as used by Zahra et al. [11] approach. The rule implemented for this category is not purely a grammar rule but a Bag-of-Words technique as follows:

RULE-2; LEMMA in PreceptualSensesWordList ()

3.4.3 Feature-3: Reporting Impact of Disaster

Reporting the impact of the disaster in tweet content is a common practice during disaster events. It can be useful in evaluating the severity of the impact. Such tweets may contain words like school canceled, and flight delayed, etc. For implementation purpose, we have proposed a rule that is not purely grammar-based; however, a combination of both the language structure and the list of the impact-words like 'cancel', 'delay', 'suspend', 'lost', 'postpone', 'defer', 'reschedule', 'rearrange'. We have not utilized a pre-generated list of events but used a rule that combines with a list of impact-words that helps in finding disaster impact features. The implementation of this feature would be as follows:

RULE-3; POS like 'NN%' and OnTarget (DisasterImpacts (LEMMA))

3.4.4 Feature-4: Words Indicating the Intensity of the Disaster

In this category of features, the feature-words show the severity of the disaster, like intense, strong, dangerous, big, etc. Such feature words can help the institutions in decision making and help the response teams to react accordingly. For implementation, we have proposed a rule as follows:

RULE-4; POS= 'JJ' and DepRel= 'nsubj' and IDX <HeadIDX and NER= 'O'

and OnTarget(POS like 'NN%' and NER= 'CAUSE_OF_DEATH')

3.4.5 Feature-5: First-Person Pronouns and Adjectives

This feature type has a standard list to follow, and its implementation is straightforward. For this purpose, we have employed Bag-of-words (BOW) technique as employed by the Zahra et al. [11] approach. The rule for this feature is implemented as follows:

RULE-5; WORD in FirstPersonProNouns()

3.4.6 Feature-6: Personalized Location Markers

This feature is also an important feature to indicate the density of disaster's impact by extracting the personalized location markers. It contains terms like my office, our area, our city, etc. It is difficult to distinguish a noun of locations from other nouns. For this purpose, we used a list of location nouns [40]. For the implementation purpose, we have proposed the rules as follows:

RULE-6; POS= 'PRP$' and DepRel= 'nmod:poss' and IDX <HeadIDX and NER= 'O' and

LEMMA in ('my','we') and OnTarget(POS= 'NN' and NER= 'O' and LocationNoun(WORD))

3.4.7 Feature-7: Exclamation and Question Marks

This feature can easily be acquired using a defined list of words that are "!" and "?" The formed rule to extract this feature is as follows:

RULE-7; WORD in ('?','!')

3.4.8 Feature-8: Expletives

Expletives are words or phrases which are not required for an explanation of the basic meaning of the sentence, e.g., wtf, omg, s**t, etc. For the implementation of this feature, slang words lists are used by Zahra et al. [11] approach, the same method is adopted in this work. It is not possible to identify them from sentences using grammar rules. The list is available online [41].

RULE-8; WORD in ListOfSlangWords()

3.4.9 Feature-9: Mention of a Routine Activity

This feature type includes the words that explain routine activates e.g., sleeping, watching a movie. Implementation of this type is proposed by using the word list available online [42]. In our work, we have proposed the following grammar rule for implementation purpose;

RULE-9; POS= 'VBG' and NER= 'O' and WORD <> 'GON'

3.4.10 Feature-10: Time Indicating Words

Time indicating words are considered as an important feature to understand and react to a disaster event. For example, now, at the moment, just, etc. For implementation purpose, we have proposed rules as follows:

RULE-10; POS= 'RB' and DepRel= 'advmod' and((WORD= 'just') or

(WORD<> 'just' and NER= 'DATE')) and OnTarget(POS in 'RB','VBP','JJ')

3.4.11 Feature-11: Short Tweet-length

This feature identification requires other features to explain the actual concept. For example, a tweet with a single word like "earthquake". In the current context the feature is true; however it does not convey the true meaning rather the word "earthquake!" changes the strength of the tweet. We have not found standard word-count in the literature therefore this feature implementation is complex. We set the tweet length to be less than or equals to 8 words for implementation purposes. This remains an open area for future work in this field. Implementation rule would be as follows;

RULE-11; WordCount([Tweet Text])

3.4.12 Feature-12: Caution and Advice for Others

This feature includes words like watch out, be careful, etc. to intimate others about an event. It is implemented using dictionaries, but we have proposed the grammar rules to identify such words from the tweet content. For implementation purpose we have proposed rules as follows;

RULE-12; POS= 'VB' and DepRel= 'cop' and IDX <HeadIDX and OnTarget(POS= 'JJ')

3.4.13 Feature-13: Mention of Disaster Locations

We used the feature of the CoreNLP tool to identify the mentioned location from tweet content. For implementation purpose, we have proposed the rules as follows:

RULE-13; LOCATION_TYPE(NER) OR WORD in ('north', 'south', 'east', 'west')

3.5 Feature Evaluation

The existence of individual features like "Exclamation and question marks", "Short tweet-length", "First-person pronouns", "Words indicating perceptual senses", do not explain the concept that the tweet is posted by a true eyewitness. In order to incorporate such situations, we have generalized the idea and adopted a standardized approach. It states that a tweet having at-least two identified features would be marked as an eyewitness tweet. A non-eyewitness feature-based [22] tweet is marked as non-eyewitness whereas tweets with no identified features are categorized as unknown. An exceptional scenario is adopted for the "Exclamation and question marks", wherein even only one feature is identified tweet is marked as "Eyewitness".

3.6 Evaluation Parameters

The proposed approach is evaluated in terms of precision, recall, and f-measure to measure the effectiveness. These evaluation matrices are commonly used in this field. Zahra et al. approach and other approaches in the literature also exploit these evaluation measures.

3.7 Evaluation Strategy

The proposed approach is evaluated against the results generated by Zahra et al. approach on a specific data set [11]. If the proposed approach generated the same result as Zahra et al. approach, then it is considered

as a true-match. In this manner values for precision and recall are calculated for each identified feature. The f-score is then calculated to have a harmonic mean of both precision and recall values.

4 Experiments and Results

This section presents the overall evaluation results of the proposed approach. The following subsections briefly explain the results generated by each step of the proposed methodology.

4.1 Dataset and Pre-Processing of Tweets

The proposed approach is evaluated using a dataset containing 8000 tweets on earthquake, flood, hurricane, and wildfire. Every tweet in the dataset contains the annotations done by the Zahra et al. approach. A brief statistics of all the categories are illustrated in Tab. 5.

Table 5: Dataset statistics (tweets)

Category	Earthquake	Flood	Hurricanes	Wildfire
Eyewitness	1600	980	791	884
Non-Eyewitness	200	679	780	880
Unknown	200	341	429	236
Total	**2000**	**2000**	**2000**	**2000**

As the dataset has already been evaluated and updated by Zahra et al. approach, it contained a small number of tweets which required pre-processing steps. The existence “ ‘ “ and comma was found in 2942 tweets, and they were updated during the pre-processing. The dataset contains the author generated results against each tweet. The same dataset is used as the benchmark to compare proposed approach results with Zahra et al. results.

4.2 Parsing and Feature Extraction

Parsing of all tweets for Part-Of-Speech (POS) tagging, Lemmatization, Named Entry, and Entity relation, Stanford CoreNLP tool is used. The content of all 8000 tweets were passed through the tool that eventually generated the results having 22,932 rows with all required fields, as discussed in Section-3.3.

After passing all tweets from the Stanford CoreNLP tool, the results are then thoroughly studied and analyzed for the language structure. The relationship among different words of a sentence was studied, and based on these findings, and rules were generated for feature extraction (as discussed in Section 3.4).

The feature-words count is not shared by the authors of the Zahra et al. approach. However proposed approach uses grammar rules to validate the accuracy of feature-words extraction. Each tweet has been annotated manually for the identification of feature-words. The dataset, along with features list, was provided to a Natural Language Processing expert with excellent language skills and has published different articles in the field of semantics analysis and information extraction. The expert is a Ph.D. Scholar in the field of semantic computing.

The proposed grammar rules described in Section 3.4 were applied to the processed dataset of 8000 tweets, and the feature-words were extracted against each grammar rule. Manually extracted feature-words by the expert for each feature using 8000 tweets were then compared with the list of automatically generated results of the proposed approach of grammar rules. The comparison results yielded an accuracy

of 0.95. The resulting dataset was then used for further calculations, evaluations, and comparison with Zahra et al. approach [11].

4.3 Results of Proposed Approach

The results were generated by the proposed approach are shown in Tab. 6.

Table 6: Proposed approach result

Disaster Type >	Earthquake			Flood			Hurricane			Wildfire		
Category	Ew	NEw	Un	Ew	NEw	Un	Ew	NEw	Un	Ew	NEw	Un
Eyewitness(Ew)	1214	330	56	356	212	412	297	327	167	119	545	220
Non-Eyewitness (NEw)	74	96	30	193	223	263	112	552	116	55	676	149
Unknown(Un)	97	78	25	78	116	147	56	320	53	15	158	63
TOTAL	**1385**	**504**	**111**	**627**	**551**	**822**	**465**	**1199**	**336**	**189**	**1379**	**432**

The proposed approach results generated for each disaster type for all evaluation parameters (precision, recall, and f-score) are illustrated in Tab. 6. To align the table, the category names are denoted as; "Ew" for eyewitness, "NEw" for non-eyewitness, and "Un" for unknown categories.

It was observed that the "Earthquake" disaster category achieved high-value results for precision, recall, and f-score. The remaining disaster categories have achieved average results for all the parameters. The same proportion of results was observed in Zahra et al. approach. Best results were generated with earthquake events as the maximum numbers of features were identified. During manual analysis, it was observed that different characteristics were found for different disaster types due to their different nature.

4.4 Comparison of Proposed Approach with Zahra et al. Approach

Each feature type was evaluated independently, and results were compared with Zahra et al. approach. Comparison of proposed approach and the Zahra et al. approach is illustrated in Fig. 4 in terms of precision, recall, and F-Scores measures. Zahra et al. approach uses a manually created static dictionary for the evaluation which is not scalable. For instance, consider the following scenarios:

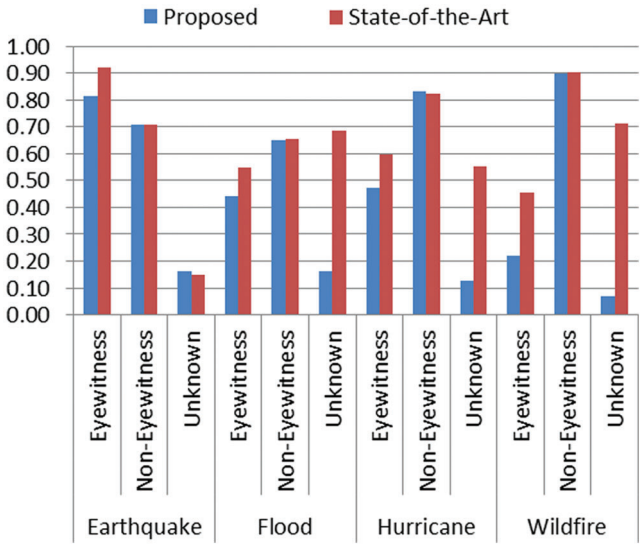

Figure 4: Comparison scores of proposed and state-of-the-art approaches

a) When the domain changes, the list of words might change
b) The unseen tweets may contain different words or vocabulary that will directly affect the performance of the state-of-the-art approach.

In the event of a disaster, it will be a cumbersome effort to process millions of tweets to identify the eyewitness in real-time as it will involve thousands of domain-exports to update the dictionary. In order to cover this gap, the language structure, linguistic features, and word relations have been critically reviewed to automatically identify the eyewitness. Subsequently, the idea of creating grammar rules for the automatic extraction of eyewitness feature-words has been used. The grammar-based rules have been created for feature extraction for all thirteen identified features without using static dictionaries with reasonable accuracy.

The proposed approach was then evaluated for all disaster-types, including earthquakes, floods, hurricanes, and fire. Based on the static dictionary, the Zahra et al. approach was able to produce an F-Score value of 0.92 for Eyewitness identification in the earthquake category. The proposed approach secured F-Score values of 0.81 in the same category. This score can be considered as a significant score without using a static dictionary. For the non-eyewitness category, Zahra et al. approach and proposed approaches both secured F-Score of 0.71, as the same approach of Doggett et al. [22] was adopted. For the unknown category, the proposed approach performed better and achieved the F-Score of 0.16, while the Zahra et al. approach achieved the F-Score of 0.15. The results for other disaster types remained in the same proportion for both approaches.

The proposed technique produced lower results for the "wildfire" disaster type in comparison to the results achieved for the remaining disaster types. Because the proposed approach automatically extracts feature-words and do not involve human interaction for manual creation and maintenance of dictionaries for feature-words. We claim that the proposed approach could be deemed as a potential approach, especially in the scenarios when we have millions of tweets to process in real-time.

5 Conclusion

Twitter is considered as a potential news beaker platform. People across the globe share information to their followers regarding incidents they witness, especially disastrous events like flood, earthquake, fire, etc. It is considered a challenging task to identify eyewitness tweets about the incident from the millions of tweets shared by twitter users.

Whereas on other hand the information shared by an eyewitness is valuable in terms of understanding the true intensity of the event. Emergency services can use such information for disaster management. The recent approach proposed by Zahra et al. [11] has identified various characteristics from the content of the tweet to identify the feature words. However, they have manually created static dictionary for the identification of feature-words. A critical analysis of literature reveals three major limitations of Zahra et al. approach for using a static dictionary. Firstly, need to be updated with the changing in the domain. Secondly, the involvement of domain-experts will be required for updating the static dictionaries. Thirdly, handling millions of tweets in real-time is daunting task.

This research paper has proposed a comprehensive methodology for automatic extraction of feature set. A fine-tuned grammar rules were created after studying the language structure. The grammatical rule for each of the feature was carefully designed by detailed analysis of the dataset and features were extracted based on these rules. These features were further utilized to categorize the tweets into different categories like "eyewitness", "non-eyewitness", or "unknown" [11].

The proposed approach results were evaluated in terms of precision, recall, and f-measure on a benchmark dataset of 8000 tweets. Furthermore, the results were also compared with the Zahra et al.

approach. The proposed approach secured F-Score of 0.81 for eyewitness tweets for earthquake incident. This score can be considered as a significant score without using a static dictionary. Hence, the proposed approach can be a potential approach considering no human effort is required in terms of dictionary management. Similarly, millions of tweets can be processed on real time processing. Last but not least a generalized approach which can work with various disaster types. The proposed approach is a novel contribution to the concept of automatic identification of eyewitness tweets using grammar rules on twitter without human interaction.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] M. Imran, C. Castillo, F. Diaz and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM Computing Surveys, CSUR*, vol. 47, no. 4, pp. 1–38, 2015.

[2] S. Vieweg, A. L. Hughes, K. Starbird and L. Palen, "Microblogging during two natural hazards events: What twitter may contribute to situational awareness," in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, Atlanta Georgia, USA, pp. 1079–1088, 2010.

[3] W3C, "*Twitter Usage Statistics*," 2020. [Online]. Available: https://www.internetlivestats.com/twitter-statistics/.

[4] H. Kwak, C. Lee, H. Park and S. Moon, "What is twitter, a social network or a news media?," in *Proc. of the 19th Int. Conf. on World Wide Web*, Raleigh North, Carolina, USA, pp. 591–600, 2010.

[5] CNA, "*'Middle of the ocean': Delta flight from Beijing makes emergency landing on remote Alaskan island*," 2008. [Online]. Available: https://www.channelnewsasia.com/news/world/delta-flight-middle-of-the-ocean-seattle-beijing-emergency-land-11062706.

[6] C. Beaumont, "*New York plane crash: Twitter breaks the news, again*," 2009. [Online]. Available: https://www.telegraph.co.uk/technology/twitter/4269765/New-York-plane-crash-Twitter-breaks-the-news-again.html.

[7] USGS, "*U.S. Geological Survey*," 2019. [Online]. Available: https://www.usgs.gov/.

[8] M. Milian, "*Twitter sees earth-shaking activity during SoCal quake*," 2008. [Online]. Available: https://latimesblogs.latimes.com/technology/2008/07/twitter-earthqu.html.

[9] Wikipedia, "*Boston Marathon bombing*," 2020. [Online]. Available: https://en.wikipedia.org/wiki/Boston_Marathon_bombing.

[10] Wikipedia, "*Westgate shopping mall attack*," 2020. [Online]. Available: https://en.wikipedia.org/wiki/Westgate_shopping_mall_attack.

[11] K. Zahra, M. Imran and F. O. Ostermann, "Automatic identification of eyewitness messages on twitter during disasters," *Information Processing & Management*, vol. 57, no. 1, pp. 1–15, 2020.

[12] C. Amaratunga, "Building community disaster resilience through a virtual community of practice (vcop)," *International Journal of Disaster Resilience in the Built Environment*, vol. 5, no. 1, pp. 66–78, 2014.

[13] Y. Kryvasheyeu, H. Chen, N. Obradovich, E. Moro, P. Van Hentenryck *et al.,* "Rapid assessment of disaster damage using social media activity," *Science Advances*, vol. 2, no. 3, pp. 1–11, 2016.

[14] C. Allen, "A resource for those preparing for and responding to natural disasters, humanitarian crises, and major healthcare emergencies," *Journal of Evidence-Based Medicine*, vol. 7, no. 4, pp. 234–237, 2014.

[15] E. Schnebele and G. Cervone, "Improving remote sensing flood assessment using volunteered geographical data," *Natural Hazards and Earth System Sciences*, vol. 13, no. 3, pp. 669–677, 2013.

[16] J. Teevan, D. Ramage and M. R. Morris, "#twittersearch: A comparison of microblog search and web search," in *Proc. of the fourth ACM Int. Conf. on Web Search and Data Mining*, Hong Kong, China, pp. 35–44, 2011.

[17] B. Haworth and E. Bruce, "A review of volunteered geographic information for disaster management," *Geography Compass*, vol. 9, no. 5, pp. 237–250, 2015.

[18] P. M. Landwehr and K. M. Carley, "Social media in disaster relief, " in *Data Mining and Knowledge Discovery for Big Data*. Springer, Berlin, Heidelberg, pp. 225–257, 2014.

[19] P. Meier, "Crisis mapping in action: How open source software and global volunteer networks are changing the world, one map at a time," *Journal of Map and Geography Libraries*, vol. 8, no. 2, pp. 89–100, 2012.

[20] F. Ostermann and L. Spinsanti, "Context analysis of volunteered geographic information from social media networks to support disaster management: A case study on forest fires," *International Journal of Information Systems for Crisis Response and Management*, vol. 4, no. 4, pp. 16–37, 2012.

[21] M. Truelove, M. Vasardani and S. Winter, "Testing a model of witness accounts in social media," in *Proc. of the 8th Workshop on Geographic Information Retrieval, GIR*, Dallas Texas, pp. 1–8, 2014.

[22] E. V. Doggett and A. Cantarero, "Identifying eyewitness news-worthy events on twitter," in *Proc. of the Fourth International Workshop on Natural Language Processing for Social Media*, Austin, Texas, USA, pp. 7–13, 2016.

[23] R. Fang, A. Nourbakhsh, X. Liu, S. Shah and Q. Li, "Witness identification in twitter," in *Proc. of the 4th International Workshop on Natural Language Processing for Social Media*, USA, pp. 65–73, 2016.

[24] H. Tanev, V. Zavarella and J. Steinberger, "Monitoring disaster impact: Detecting micro-events and eyewitness reports in mainstream and social media, " in *ISCRAM*, Albi, France, pp. 592–602, 2017.

[25] S. Kumar, G. Barbier, M. A. Abbasi and H. Liu, "Tweettracker: An analysis tool for humanitarian and disaster relief," in *Fifth Int. AAAI Conf. on Weblogs and Social Media*, Barcelona, Spain, pp. 78–82, 2011.

[26] F. Abel, C. Hauff, G. J. Houben, R. Stronkman and K. Tao, "Twitcident: Fighting fire with information from social web streams," in *Proc. of the 21st Int. Conf. on WWW*, Lyon, France, pp. 305–308, 2012.

[27] M. Imran, C. Castillo, J. Lucas, P. Meier and S. Vieweg, "Aidr: Artificial intelligence for disaster response," in *Proc. of the 23rd Int. Conf. on World Wide Web*, Seoul, Republic of Korea, pp. 159–162, 2014.

[28] D. Thom, R. Krüger, T. Ertl, U. Bechstedt, A. Platz *et al.,* "Can twitter really save your life? A case study of visual social media analytics for situation awareness," in *IEEE Pacific Visualization Sym. (Pacific Vis)*, Hangzhou, China, IEEE, pp. 183–190, 2015.

[29] M. F. Goodchild and L. Li, "Assuring the quality of volunteered geographic information," *Spatial Statistics*, vol. 1, pp. 110–120, 2012.

[30] M. Truelove, M. Vasardani and S. Winter, "Towards credibility of microblogs: Characterizing witness accounts," *GeoJournal*, vol. 80, no. 3, pp. 339–359, 2015.

[31] N. Diakopoulos, M. De Choudhury and M. Naaman, "Finding and assessing social media information sources in the context of journalism," in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, Austin, Texas, USA, pp. 2451–2460, 2012.

[32] A. Olteanu, S. Vieweg and C. Castillo, "What to expect when the unexpected happens: Social media communications across crises," in *Proc. of the 18th ACM Conf. on Computer Supported Cooperative Work and Social Computing*, Vancouver, British Columbia, Canada, pp. 994–1009, 2015.

[33] S. Kumar, F. Morstatter, R. Zafarani and H. Liu, "Whom should I follow?: Identifying relevant users during crises," in *Proc. of the 24th ACM Conf. on Hypertext and Social Media*, Paris, France, pp. 139–147, 2013.

[34] F. Morstatter, N. Lubold, H. Pon-Barry, J. Pfeffer and H. Liu, "Finding eyewitness tweets during crises," *ArXivPreprint, ArXiv*, vol. 1, pp. 1403–1773, 2014.

[35] LIWC, "*Linguistic Inquiry and Word Count*," 2019. [Online]. Available: http://www.liwc.net/.

[36] J. Callicott, "*OpenCalais API*," 2019. [Online]. Available: https://www.drupal.org/project/opencalais_api.

[37] Stanford, "*Stanford CoreNLP–Natural language software*," 2019. [Online]. Available: https://stanfordnlp.github.io/CoreNLP/.

[38] K. Zahra, F. O. Ostermann and R. S. Purves, "Geographic variability of twitter usage characteristics during disaster events," *Geo-Spatial Information Science*, vol. 20, no. 3, pp. 231–240, 2017.

[39] M. C. de Marneffe and C. D. Manning, "Stanford typed dependencies manual," *Technical Report*, Stanford University, pp. 338–345, 2008, Revised 2016.

[40] WordHippo, "*What is the noun for locate?*," 2020. [Online]. Available: https://www.wordhippo.com/what-is/the-noun-for/locate.html.

[41] Annemarie, "*Your ultimate list of English internet slang on social media*," 2017. [Online]. Available: https://www.speakconfidentenglish.com/english-internet-slang.

[42] Woodward, "*Daily routines English Vocabulary*," 2020. [Online]. Available: https://www.vocabulary.cl/Lists/Daily_Routines.htm.