

Analyzing the Sentiment of Crowd for Improving the Emergency Response Services

Neha Singh, Nirmalya Roy, Aryya Gangopadhyay

Department of Information Systems

University of Maryland Baltimore County

nehasingh1@umbc.edu, nroy@umbc.edu, gangopad@umbc.edu

Abstract—Twitter is an extremely popular micro-blogging social platform with millions of users, generating thousands of tweets per second. The huge amount of Twitter data inspire the researchers to explore the trending topics, event detection and event tracking which help to postulate the fine-grained details and situation awareness. Obtaining situational awareness of any event is crucial in various application domains such as natural calamities, man made disaster and emergency responses. In this paper, we advocate that data analytics on Twitter feeds can help improve the planning and rescue operations and services as provided by the emergency personnel in the event of unusual circumstances. We take a different approach and focus on the users' emotions, concerns and feelings expressed in tweets during the emergency situations, and analyze those feelings and perceptions in the community involved during the events to provide appropriate feedback to emergency responders and local authorities. We employ sentiment analysis and change point detection techniques to process, discover and infer the spatiotemporal sentiments of the users. We analyze the tweets from recent Las Vegas shooting (Oct. 2017) and note that the changes in the polarity of the sentiments and articulation of the emotional expressions, if captured successfully can be employed as an informative tool for providing feedback to EMS.

Index Terms—Twitter, Sentiment Analysis, Emotion Detection, Emergency services, Change Point Detection

I. INTRODUCTION

Twitter has been there for more than a decade as a medium that allow millions of Twitter users to utilize this social platform worldwide to share the information, opinion, and broadcasting the short messages [1]. Twitter users are also known as social sensors [2], [3], since they provide the information about the event as soon as it happens and eventually the discussion starts percolating on as the event unfolds. There are many applications where twitter data has been leveraged successfully in a collaborative environment [4], [5]. Twitter provides the gigantic amount of data about any topic that made it useful tool for the researchers to exploit this data to identify various task such as event detection [6], event tracking [7], [8], event monitoring [9] etc. Sentiment Analysis (SA) [10] is another domain which primarily focuses from the users' point of view and their sentiments about the various aspects of the event, services, product or any topic. The Sentiment Analysis summarizes the users' emotions and opinions about different topics and gives rise an aggregated idea about the user's perception and insights which is not covered in event oriented analysis or event driven task. Sentiment Analysis and Opinion Mining is a growing field of research that uses

Natural Language Processing [11], Information Retrieval [12], text analysis [13] and computational linguistics [14] to extract user's opinions and emotions towards any topics or event [15], [16], [17], [18]. Initially, sentiment analysis was used as the voice of consumers that can identify the customer's sentiments towards any services or product. This analysis was extremely helpful for the businesses and organizations because it can provide the constructive feedback to the businesses, brands and organizations so that they can improve their services. Recently, sentiment analysis has also been applied during emergency events that deals with the sentiment's polarity (positive, negative or neutral) and various emotions of the users during the emergency event [19]. The critical details about the emotions and concerns of the people towards the crisis or emergency event are significantly important for the emergency responders to consider and reflect for improving and better managing their services to the community. Humans are driven by the sentiments most of the time, therefore, the current or future help tuned to that perception of the public will be very effective. For example, if people are fearful and angry because of some events, they are most likely not ready to listen indeed if local authorities and emergency responders have better idea about the general concern and panic of the people, it will be much easier to manage the crowd. Motivated by this, in this paper we take a bold step to combine event detection after sentiment analysis on twitter to mine the fine-grained perception of the crowd which can be piggybacked to the emergency personnel in the form of reviews of their services provided in critical situation. We investigate change point detection algorithm to detect the changing trend in the emotions and identify the associated tweets and events that caused the change in the emotion in order to send the feedback to the first responders to deal with the situation more emphatically. The main contributions of this paper can be summarized as follows:

- We collect the tweets from Las Vegas shooting event and process this dataset for inferring the sentiments and emotions of the crowd.
- We localize the various events with topic model to categorize them in their respective situational awareness.
- We propose a Bayesian change point detection technique to identify the changes and trends in emotions over time to provide real-time feedback to the local authorities.

Section 2 of the paper discusses the related work and Section 3 presents the data processing methodology associating with sentiment analysis and emotion detection on the Las Vegas shooting event. Section 4 describes the overall process of the methodology. Section 5 discusses about the results and Section 6 articulates the challenges and limitation of this work. Section 7 briefly explains the future goals of this research followed by the conclusion.

II. RELATED WORK

Sentiment analysis (SA) and opinion mining are growing field of research that use Natural Language Processing (NLP) and Information Retrieval (IR) to extract the users' comments, reviews, emotions and opinion towards the product, topic or event [20], [21]. The recent trends in SA, and its extended application to other areas like transfer learning, emotion detection, building resources have been mentioned in [21]. Classification process can be performed mainly on three levels such as aspect levels, sentence levels and document levels. Various machine learning and Lexicon-based approaches have been used for classification of the polarity of sentiments with Naive Bayes, SVM and WordNet being the most common among all. [19] has focused on identifying individual's sentiments polarity during natural disaster. It has analyzed the tweets during Hurricane Sandy and projected the sentiments of users on a geographical map. The analysis shows that user's sentiments change based on their location and distance from the disaster events which may be helpful for emergency responders to get stronger situational awareness. [22] applied sentiment analysis on tweets of Genoa flood happened in October 2014 and attempted to identify the useful tweets that can provide information from the disaster management perspective. Subjective, negative, positive or ironic were the main labels for the sentiments in their dataset. The research shows that negative sentiments were more insightful than positive ones and there are specific situations that could be easily missed in complete dataset. [23] explores the usage of sentiment mining and exploits the social media in disaster management to detect people reaction during crisis situation enabling the better situation awareness. SentiStory [24] proposes a system that summarizes the events after eliminating the redundancy in microblogs based on coarse-grained sentiment analysis and detecting the significant changes and causes of changes in the microblogs. This Research [25] present a technique for visualizing the emotional state of the people during different disaster event through the analysis of tweets. This recent study [26] has showed that change point detection is useful in identifying potential sub-events that causes the changes in sentiments of public over Twitter using the Delhi Election 2015 tweets data. The authors have used the average sentiment score (positive and negative) and applied the change point detection algorithm to see the significant changes but in this research we are focus more on identifying the temporal changes in every emotion. [19] has used the sentiment labels as positive, negative and neutral along with geo-tagged tweets but in this work we are enhancing the situation awareness by

adding more fine-grained contexts through additional emotions such as anger, joy, sadness etc. SentiStory [24] posited fine-grained emotions but we propose to extract the eight different kind of emotions and capture the changes with the emotional context. In this work, we first localize a single event using a topic model and user's emotions through tweets during and aftermath of the event. In contrast to [26], where only the positive and negative sentiments of the public have been investigated, we propose to look at the finer categorization of the people sentiments such as anger, anticipation, disgust, fear, joy, sadness, surprise and trust. We posit that these micro-level emotion type detection from the crowd's aggregated social footprints are insightful to design future EMS services for the population in accessing and addressing the event situation with better emotional understanding.

III. DATA PROCESSING FOR SENTIMENT ANALYSIS APPROACH

In this section, we describe the necessary steps for processing the Las Vegas shooting dataset for sentiment analysis.

A. Data Description and Data Collection

The data set used for this study is collected from Twitter right after the shooting happened in Las Vegas on October 1st, 2017 using the Twitter REST API. The data was collected using the keywords such as “#LasVegas”, “#LasVegas” and “#LasVegasShooting” from October 1st, 2017 till October 6th, 2017. This tragic event left many lives lost and hundreds of people injured. This event quickly got the attention of people and Twitter users shared their opinion and emotions about this horrendous event. The data that we collected has total of 119487 tweets in total and after removing the duplicate tweets we left with 34859 tweets. The sentiment analysis is performed on the emotions of the users but the data also include some spam and advertisement tweets on a regular interval about their business and services which is not the target of this analysis. Thus, we removed those tweets posted by those twitter handles. Finally, 32909 tweets were taken for further analysis along with other data fields such as tweet id, created, retweet, retweetCount etc.

B. Data Preprocessing

The data is collected using R package “*twitteR*” and data analysis with the help of “*tm*” packages in R. The shooting event happened on October 1, 2017 and the number of tweet distribution per day can be seen in Fig. 1(a), the primary motivation of this analysis is to understand the emotions of the people during and after the event happened. The number of tweets varies according to the intensity of the event and the time of the day. In general afternoon and late night time have the highest intensity in number of tweets. Furthermore, there are some steps taken for cleaning of the data which we depict next.

C. Data Cleaning and Visualization

The data collected and saved in the csv format which contains tweets along with retweets, numbers, url's, white spaces, special characters, user mentions, hashtags and punctuation. In data cleaning all the words are transformed into lower letters followed by removing the stopwords. Fig. 1(b) shows the highest frequent terms with the respective frequency and the top five most frequent words are "shooting", "victims", "gun", "people". The words "LasVegas", "US" and "vegas" were removed from the data, since the data was collected using these keywords. Data cleaning is very important for the visualization purposes but in sentiment analysis some aspects like hashtags, punctuation, emoji's etc., are crucial and gives a lot of information in correctly identifying the emotions, thus we keep them according to the application requirements.

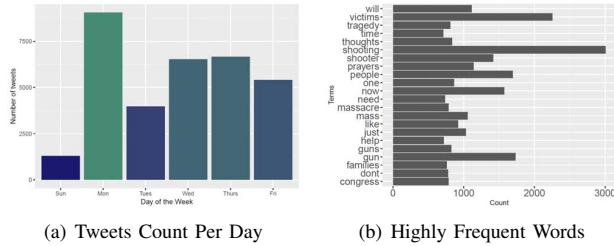


Fig. 1. Data Description

IV. OVERALL PROCESS

In this section, we describe the step by step process of our twitter data analysis for sentiment categorization, and feedback mapping. Fig. 2 depicts the entire process. First, it starts with collecting twitter data and cleaning it for better visualization and understanding the characteristics of the data. Second, data is cleaned and processed in the required form where it can be analyzed further. Third, Sentiment analysis has been employed in order to extract emotions out of data that identify eight different emotions where perceptions of the victims are of utmost importance. Fourth, Bayesian change point detection algorithm is used to identify the changes in the emotions. Finally, these identified changes points represent the combination of tweets that summarize the cause of the change in emotion thus reporting it to first responders to help them better understand the general public sentiments during the crisis.

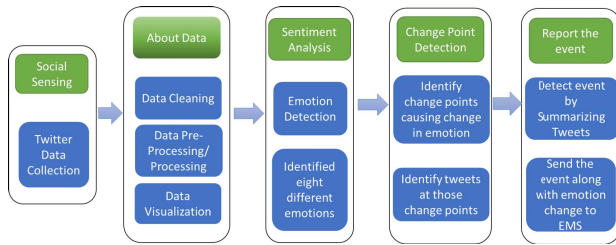


Fig. 2. Process flow chart

A. Topic Modeling for Event Detection

Latent Dirichlet Allocation (LDA) is the simplest topic modeling, which is a traditional approach to detects main theme in a document and it has been applied to various kinds of data [27]. We apply the LDA in our dataset to detect the overall daily theme in the data, thus dividing our data on daily basis and perform the simple topic modeling with Gibbs sampling and 1000 iterations in order to identify the major events happened that day. The results of the topic modeling are shown below in the tables representing topics from each day. The key observation here is that from day 1 and day 2, the topic modeling does not provide us any word related to shooting or about the shooting event. However, day 3 onward does provide us with the topics such as "shooting", "gun", "victims", "killed", "prayers" etc. which are clearly related to the shooting event, day wise result of topic modeling shown in table I, II and III for the respective date. Topic modeling does not help us in finding out the related keywords and topics to the shooting event from the day one. The emergency first responders need the information right from the very beginning of the event/day in order to rescue or provide appropriate help to the public. Thus, Topic modeling alone is not sufficient for the first responders, we need to include the sentiment analysis and change point detection algorithm in order to provide the timely and correct information about the event.

TABLE I
TOP FIVE TOPICS AND RESPECTIVE WORDS FOR DAY 1 (OCTOBER 1ST)
USING LDA

Topics 1	Topics 2	Topics 3	Topics 4	Topics 5
great	free	see	vegas	the
oct	club	more	las	prison
new	strip	just	now	years
need	pick	good	night	home
real	call	want	today	latest
this	credit	memories	were	thanks
humidity	limo	entrepreneur	video	here
temperature	have	tomorrowucufef	saturday	lasvegassun
kmh	live	cant	you	morning
team	code	time	will	day

TABLE II
TOP FIVE TOPICS AND RESPECTIVE WORDS FOR DAY 2 (OCTOBER 2ND)
USING LDA

Topics 1	Topics 2	Topics 3	Topics 4	Topics 5
see	vegas	free	more	the
prison	las	club	just	great
good	now	strip	oct	new
years	home	pick	want	get
live	this	call	today	night
tomorrowucufef	need	limo	october	entrepreneur
memories	real	have	cant	time
your	you	credit	help	one
here	latest	code	humidity	its
morning	were	recommend	play	lasvegassun

B. Sentiment and Emotion Analysis over time

The sentiment of the data is analyzed with the help of R package "sentiment" by Timothy Jurka. This package is

TABLE III
TOP FIVE TOPICS AND RESPECTIVE WORDS FOR DAY 3 (OCTOBER 3RD)
USING LDA

Topics 1	Topics 2	Topics 3	Topics 4	Topics 5
victims	shooter	shooting	gun	now
vegas	attack	the	just	congress
las	massacre	will	need	you re
prayers	people	mass	like	weapons
families	hotel	news	one	ban
help	what	latest	people	assault
thoughts	room	killed	dont	demand
our	paddock	trump	how	outraged
tragedy	many	pray	time	please
love	shooters	violence	this	blood

utilized mainly for classifying the sentiment polarity i.e. positive, negative and neutral. Naive bayes classifier is used for this classification of polarity and it is trained on Janyce Wiebes subjectivity lexicon. Fig. 4(a) shows the total sentiment polarity with time in the data which is helpful in determining the overall sentiments' polarity changing with time. The negative sentiment is increasing just after the terrible shooting occurred and as the news spread more users were feeling bad and condemning the act which is the result of negative sentiments in users tweets. Though it is clear that negative sentiments are greater than the positive but it is still not helpful in forming any conclusive statements about the exact public sentiments and opinions. The negative sentiments could indicate various things including anger and fear but detailed emotions would be much more helpful. The other R package which is used to classify different emotions is "syuzhet" package developed in the NLP group at Stanford. This package calls the NRC Word-Emotion Association Lexicon to compute the presence of eight different type of emotions ("anger", "anticipation", "disgust", "fear", "joy", "sadness", "surprise", "trust") and their corresponding valence in the input text file. This Lexicon was prepared by choosing 10,000 words from previous lexicons and labeled by Amazon Mechanical Turk to build this NRC word-association lexicon. This is based on unigrams and association scores works in binary mode (yes/no). We used this lexicon based approach to classify our data into eight different emotions and Fig. 3 shows the total emotions present in the data with the two sub images Fig. 3(a) shows the bar graph with the different emotions showing the distribution of the emotions exist in the data which gives much clearer picture of the emotion of the public about the event. Fig. 3(b) gives the amount of emotion in terms of percentage in the data. It shows that fearful emotions is the maximum in the data which can be understood as the shooting was a terrible event.

Furthermore, the plot in Fig. 4(b) shows the different emotion changing with the time. This time plot gives the crucial understanding of different emotions type and their respective intensity with time. The emotions' intensity can increase or decrease overtime but when this is happening

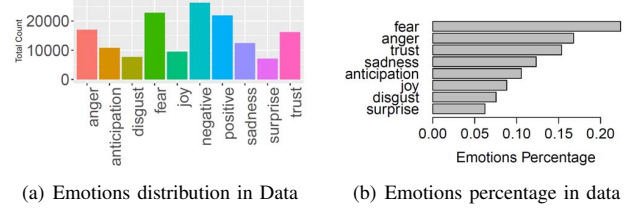


Fig. 3. Total Emotions in the Data

could be significantly important to the local officials and emergency responders at that time. The changing nature of these emotions can be linked to the services provided that time and news available during the event. These links can worked as the constructive feedback to the governments and other responsible agencies to work in more effective manners. These different emotions can also be interpreted by the officials that people liked the service during and post event or there could be more improvements towards it.

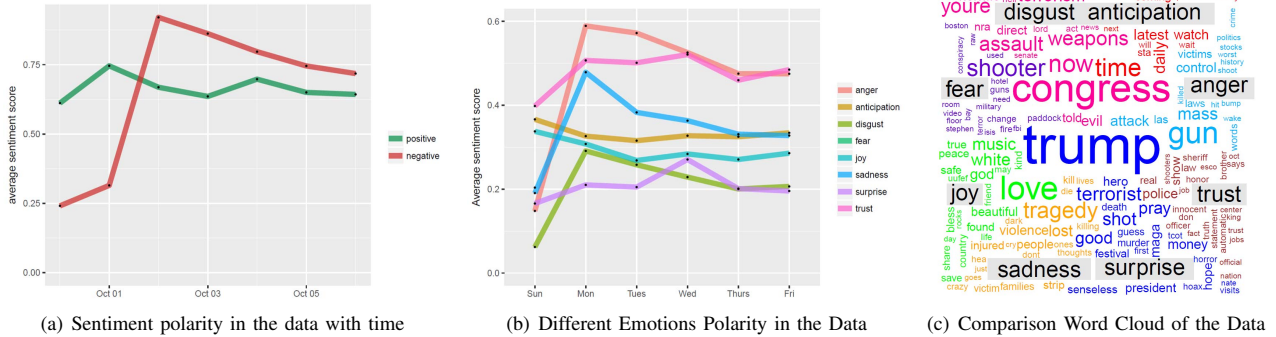
The comparison cloud is generated to have a deeper understanding seeing the most frequent words that fall into each emotions type. The Fig. 4(c) shows the comparison cloud with all the emotion types. Also, different emotions captures different words and many words fall into two or more emotions resulting in overlapping tweets with other emotions. There are also some cases in which repetition of some words exist in different emotions but they might have used in different context. For example word "Shooting" used in Angry and Fear emotion can be seen in table IV and table V.

TABLE IV
ANGRY TWEETS

No.	Angry Tweets
1	UPDATE Statement from VP fired for shameful comments about shooting victims
2	Death toll in makes it the deadliest shooting in US history Other major shootings Orlando Virginia Tech
3	Horried to hear about the shooting in My thoughts are with the victims and their families Praying for everyone
4	I get sick watching the videos of the shooting in horrible disgusting amp heartbreaking
5	Looks like somebody is shooting at the crowds with a machine gun randomly

TABLE V
FEAR TWEETS

No.	Fear Tweets
1	Churches help survivors prepare for longterm needs of more than people traumatized by shooting
2	killed amp injured in shooting Thoughts n prayers arent enoughWorld MUST stand up against terror
3	BreakingWoman told group of concert goers youre all going to die tonight minutes before shootinght
4	In light of the shooting here are ways to stay safe on amp off campus at Captioned
5	We need your help Anyone with photos or videos from shooting please contact our partners CALLFBI



We also categorized the tweets based on their emotions to understand the overlapping of tweets in different emotions and the different distributions of tweets in each emotion category resulting as fear having 8000 tweets, angry having 12000 tweets, anticipation having 8000 tweets, disgust having 6000 tweets, joy having 6000 tweets, sad having 9000 tweets, surprise having 5000 tweets and trust having 11000 tweets approximately. It seems there are a lot of overlapping among some categories but we will address this in our future work.

C. Change Point Detection in Various Emotions

As it can be seen very clearly from the Fig. 4(b) that emotions are changing with time and we are extremely interested in capturing those changes. The changes present in the sentiments can help in detecting the significant change points using the Bayesian change point detection (BCPD) algorithm introduced by Barry and Hartigan in 1993 [28] that uses the product partition model for the normal errors change point problem using Markov Chain Monte Carlo. We used the R package for the BCPD [29] to find various change points in the input sequence. The change detection for sentiment can be represented as the univariate Change Point Detection problem in a Time Series Data.

Let $x(t)$ belongs to input data X (a single dimensional data generated at time t). To formulate the problem of change point detection in a time series data by assuming that $y(t)$ is a single dimensional time series instance of data generated at the time t . The algorithm uses a unknown partition $\rho = (U_1, U_2, \dots, U_n)$ into continuous blocks, where $U_i = 1$ indicates a change point at position $i+1$ and the transition between blocks are known as “*change points*”. It initializes U_i to 0 for the positions $i < n$, with $U_n \equiv 1$. A value of U_i is computed from the conditional distribution of U_i , from the data and the current partition in every step of the Markov chain at each position i . Also, p as the probability of a change point at position i , independently for all i [29].

The change point detection algorithm assumes that the input observations X_i are independent $N(\mu_i, \sigma^2)$, where this notation specifies that μ_i are equal for all i within a block. The prior distribution of $\mu_{i,j}$, which means that the mean of the block beginning at position $i + 1$ and ending at position

j , is chosen as $N(\mu_0, \sigma_0^2/(j-i))$. One key observation in this process is that larger deviations from μ_0 are expected when there are shorter blocks, but weak signals can also be detected when adequate amount of data are available. w is defined as $\frac{\mu^2}{\mu^2 + \mu_0^2}$ for convenience and $\mu_i \equiv \mu_{ij}$ for all i in block ij to avoid confusion. The transition probability, p , for the conditional probability of a change point at the position $i+1$, may be obtained from the simplified ratio presented by Barry and Hartigan

$$\frac{p_i}{1-p_i} = \frac{P(U_i = 1|X, U_j, j \neq i)}{P(U_i = 0|X, U_j, j \neq i)}$$

$$= \frac{\left[\int_0^\gamma p^b (1-p)^{n-b-1} dp \right] \left[\int_0^\lambda \frac{w^{b/2}}{(W_1 + B_1 w)^{(n-1)/2}} dw \right]}{\left[\int_0^\gamma p^{b-1} (1-p)^{n-b} dp \right] \left[\int_0^\lambda \frac{w^{(b-1)/2}}{(W_0 + B_0 w)^{(n-1)/2}} dw \right]}$$

Here, b denotes the number of blocks obtained if $U_i = 0$, conditional on U_j , for $i \neq j$, W_0, B_0, W_1 and B_1 are the within and between block sums of squares obtained when $U_i = 0$ and $U_i = 1$ respectively, and X is the data. The tuning parameters γ and λ may take values in $[0, 1]$. The calculation of $f(U_i = 1|X, U_j, j \neq i)$ is done and sample of U_i taken at random. Once pass is completed of the data X the calculation of posterior mean μ_i at a given current position p_i is calculated using the formula

$$\hat{\mu}_r = (1 - W)\hat{X}_{ij} + x\mu_0$$

The tuning parameters μ and λ may take values in the range $[0 : 1]$ and takes care the variation in the data thus, selected such that this method is stable and effective. After each iteration, the posterior means are updated conditional on the current partition. At last, after M passes the average of the M estimates of p_i is done by approximating the posterior mean p [28].

This recent study [26] has also shown that change point detection is useful in identifying potential sub-events that causes the changes in sentiments of general public over Twitter using the Delhi Election 2015 tweets data. The research has used the average sentiment score (positive and negative) and applied the change point detection algorithm to see the

significant changes but in this research we are more focused in identifying the changes in every emotion.

The Algorithm 1 describes the complete procedure step by step. It takes the input as the raw tweets (T) with timestamp but these timestamp are not always equidistant here and returns the locations of the change points (L) as output. Once we get the locations of the change points, the tweets are extracted from those change points location. Summarize the tweets with some more surrounding tweets to get the clear understanding of the event that caused the change in the emotion. These events along with the emotional context can be send to the EMS for better situation awareness in the crisis.

Algorithm 1 Emotional Change Detection Algorithm

```

procedure CHANGINGEMOTION (
    Input: Raw Tweets ( $T = (t_1, t_2, \dots, t_n)$ ),
    Output: Change Locations ( $L$ ) )
    Data Processing
    for All tweets  $t_i \in T$  do
         $e_i = \text{GetEmotions}(t_i)$ 
    for All emotions  $e_i \in E$  do
        BCP( $E$ )
         $\lambda$  and  $\gamma$ (tuning parameters) as 0.2
        Posterior Mean  $\mu_i$  position  $p_i$ 
        Transition Probability  $p = \frac{p_i}{1-p_i}$ 
        if  $p \geq \text{threshold}$  then
            Change in emotion = True
            Location ( $L$ ) = Extract Location of changes( $L_j$ )
            Return Location  $L$ 
        Change Detected Tweets( $T_{new}$ ) = Extract tweets from
        those location  $L$ 
        Summarize Tweets to get the event  $EVENT_{summary}$ 

```

V. RESULTS AND DISCUSSIONS

In our research each emotion has a unique role in the event so by averaging the score of respective emotion, we would be loosing some important details about the emotions thus keeping all the emotion score we apply the BCPD on each emotion. Some results are shown which include input as the sequence of emotion score for each emotion and output has the values of posterior means and the posterior probability in the plot Fig. 5. The eight plots shown in Fig. 5 and each plot provides two plots that summarize the analysis of the Bayesian change point detection. The first (upper) part of the plot shows the Posterior Means that displays the data along with the posterior mean of each location in the number of tweets. The second (lower) plot displays Posterior Probability of a Change which shows the proportion of iterations that results in a change point at each location in the number of tweets.

The spikes in the Posterior Probability part of the plot which are above certain threshold (say 0.4) can be considered as strong change points detected by the algorithm and need appropriate attention in respective emotion category. Those locations of the change points and respective tweets along with

few tweets above and below it are extracted from the original data to get the fair idea of which tweets are causing the change in emotions and can be summarized and sent to Decision makers along with the time stamp. Furthermore, tweets extracted from the location can be saved and summarized to address those issue in future emergency situations as well in order to handle them in more sophisticated way. We have controlled the threshold value using the optional parameters of Bayesian change point detection algorithm γ and λ by keeping it as 0.4, which is generally kept as 0.2 for better results but it can be varied dynamically depending on the granularity and how dense or sparse the change points are.

In Fig. 6, the change point detection for anger is expressed in terms of events that might be causing the changes with the emotion. In this process we manually checked from the ground truth information, the tweets at the change points detected from the algorithm representing the respective sentiments towards the event and the increase and decrease in the sentiments are captured nicely with the change point detection algorithm. In anger the first change point detected when many people posted about the O J Simpson (armed robber and kidnapper) got out from jail on parole after 9 years thus many people expressed their anger and fear towards the event. The second and third change points are mainly capturing the events related to dead and injures people because of shooting, the very high spike shows that people were very angry at that point because of the news of killing and injures victims all over the place. The last two change points in anger shows the events related to the gun violence and tragedy affecting people's life and no strong action taken from the president Trump.

Another emotion fear is shown in Fig. 7 with events that could possibly cause the change in the emotion fear. The first few change detection are about the same as anger one, O J Simpson free from prison, there were news and many people re-tweeted about the news and fearful about the result of his coming out of jail on parole. The next change detection in the same emotion is about the people expressing their feeling about the killings and other victims and their families, praying for each other, getting the news about the number of people killed and injured in the shooting. The ERS personnel can take advantage of this smart service system by looking at activities, experience and summary of this prior emergency event.

VI. CHALLENGES AND LIMITATION

In this section we briefly summarize the underpinning challenges we faced.

1) *Challenges using Emotion Detection:* In this study we are not utilizing all the emotions that are detected through sentiment analysis. The reason behind this, every event or topic generally have some dominating emotions involved, for example if the event is natural disaster or crisis people tend to express the emotions such as sadness, anger, fear etc., more often than joy and happiness. If the event is political or social issue, we see all kinds of mixed opinion and emotions. Thus, the event type greatly influence the kind of emotions might be produced by the public. In this study, we are using the NCR

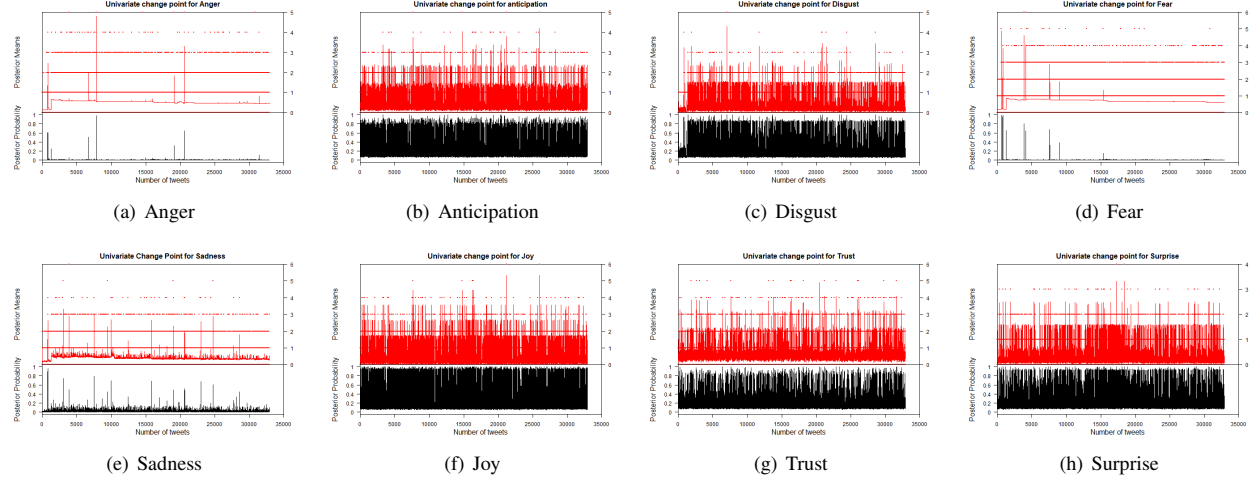


Fig. 5. Change Point Detection for Different Emotions

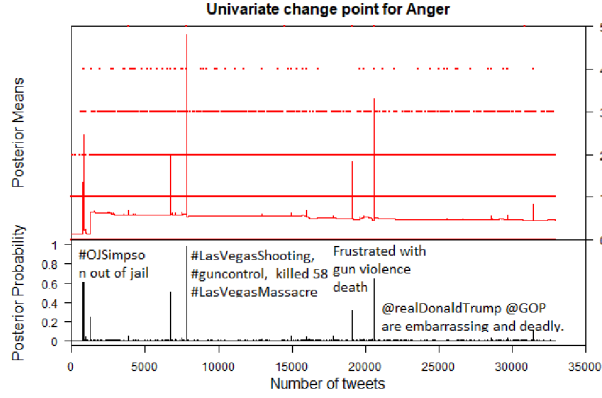


Fig. 6. Events indicating change in emotion anger

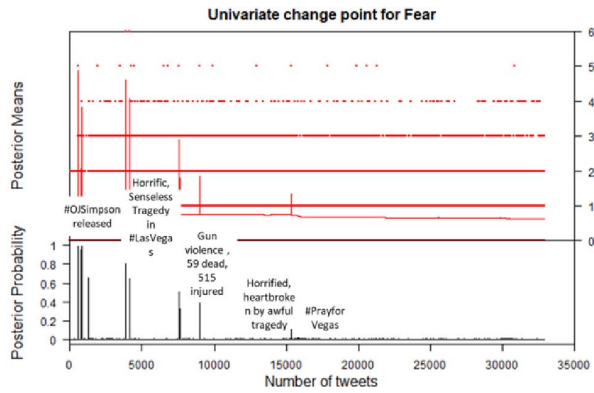


Fig. 7. Events indicating change in emotion fear

lexicon to get the eight different emotions but there are many other lexicon present to compare which lexicon will be suitable for such events. The evaluation of this dictionary based method could not be compared with some machine learning based techniques mainly because of manual annotations used therein.

2) *Challenges using Change Detection with Different Emotions:* As we can see from the Fig. 5 that change detection in the emotion Anger and emotion Fear is quite sparse in comparison to the other emotions while other emotions have very dense representation of change point. Based on the information gathered from Fig. 4(b), we can say that the emotion Anger and Fear have the similar kind of representation thus showing a significant change at specific location while other emotions have different representation, showing numerous change points at multiple locations. Also, Emotion Joy, Disgust and Anticipation have a very dense representation of change points, making it difficult to figure out the actual changes. The data distribution of the emotions is generally sparse which makes the change point detection algorithm perform really bad for some emotions and good for the dominating emotions.

VII. FUTURE WORK

In this work we have used the NCR lexicon based approach but we would also experiment with the machine learning based approach and specifically with more data, deep learning approaches to get the more accurate and robust model. We would like to compare this approach with other lexicon based sentiment analysis model to find the key similarities. Also, address the correlation among all the emotion and access the impact of it. We would like to apply the change point detection algorithm on different kind of emergency event and some more dataset to get the insights of the event through this process.

VIII. CONCLUSION

In this study, a small dataset of Las Vegas Shooting happened in October 2017, is collected and visualized. Data is analyzed based on the number of tweets collected per day

and the sentiments involved in the data. Sentiment polarity (positive and negative) is identified in the data but it is not good enough to understand the detailed emotions of the public during the crisis event. Furthermore, eight different types of emotions are identified in the dataset to get the deeper insights of the event that can be helpful for the local authorities and emergency responders to make the event handling more emotional aware. The emotions detection and tracking the changes of these emotions overtime will be more meaningful rather than tracking the event alone. Bayesian Change point detection algorithm is applied on the eight different emotions in order to identify the changes in the emotions. The locations of the change point detection are identified and the respective location and tweets are extracted to identify the cause of the change in emotions. Once the change point tweets and their surrounded tweets are summarized, it gives the context of the change in emotions and the events that causes the changes. Emotions/sentiment change along with the context of the emotional change can be identified and understood by the authorities to get the better situation awareness and deal the crisis in enlightened way.

ACKNOWLEDGMENT

This work is partially supported by the NSF CNS grant 1640625

REFERENCES

- [1] H. Kwak, C. Lee, H. Park, and S. B. Moon, "What is twitter, a social network or a news media?" in *Proceedings of the 19th International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010*, 2010, pp. 591–600.
- [2] A. Crooks, A. Croitoru, A. Stefanidis, and J. Radzikowski, "#earthquake: Twitter as a distributed sensor system," *Trans. GIS*, vol. 17, no. 1, pp. 124–147, 2013.
- [3] D. Wang, M. T. A. Amin, S. Li, T. F. Abdelzaher, L. M. Kaplan, S. Gu, C. Pan, H. Liu, C. C. Aggarwal, R. K. Ganti, X. Wang, P. Mohapatra, B. K. Szymanski, and H. K. Le, "Using humans as sensors: an estimation-theoretic perspective," in *IPSN'14, Proceedings of the 13th International Symposium on Information Processing in Sensor Networks (part of CPS Week), April 15-17, 2014, Berlin, Germany*, 2014, pp. 35–46.
- [4] M. Kranz, L. Roalter, and F. Michahelles, "Things that twitter: social networks and the internet of things," in *What can the Internet of Things do for the Citizen (CIoT) Workshop at The Eighth International Conference on Pervasive Computing (Pervasive 2010)*, 2010, pp. 1–10.
- [5] G. Sagl, B. Resch, B. Hawelka, and E. Beinat, "From social sensor data to collective human behaviour patterns: Analysing and visualising spatio-temporal dynamics in urban environments," in *Proceedings of the GI-Forum*, 2012, pp. 54–63.
- [6] F. Atefeh and W. Khreich, "A survey of techniques for event detection in twitter," *Computational Intelligence*, vol. 31, no. 1, pp. 132–164, 2015.
- [7] S. Wang, P. Giridhar, L. Kaplan, and T. Abdelzaher, "Unsupervised event tracking by integrating twitter and instagram," in *Proceedings of the 2nd International Workshop on Social Sensing*. ACM, 2017, pp. 81–86.
- [8] P. Giridhar, S. Wang, and T. Abdelzaher, "Event tracker using social networks: Demo abstract," in *Proceedings of the Second International Conference on Internet-of-Things Design and Implementation*. ACM, 2017, pp. 285–286.
- [9] P. Giridhar, T. Abdelzaher, J. George, and L. Kaplan, "On quality of event localization from social network feeds," in *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on*. IEEE, 2015, pp. 75–80.
- [10] B. Pang, L. Lee *et al.*, "Opinion mining and sentiment analysis," *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.
- [11] D. Jurafsky, "Speech and language processing: An introduction to natural language processing," *Computational linguistics, and speech recognition*, 2000.
- [12] C.-H. Chang, M. Kayed, M. R. Girgis, and K. F. Shaalan, "A survey of web information extraction systems," *IEEE transactions on knowledge and data engineering*, vol. 18, no. 10, pp. 1411–1428, 2006.
- [13] C. Nord, *Text analysis in translation: Theory, methodology, and didactic application of a model for translation-oriented text analysis*. Rodopi, 2005, no. 94.
- [14] K. W. Church and R. L. Mercer, "Introduction to the special issue on computational linguistics using large corpora," *Computational linguistics*, vol. 19, no. 1, pp. 1–24, 1993.
- [15] B. Pang and L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," in *Proceedings of the 42nd annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2004, p. 271.
- [16] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining," in *LREC*, vol. 10, no. 2010, 2010, pp. 2200–2204.
- [17] M. Etter, E. Colleoni, L. Illia, K. Meggiorin, and A. DEugenio, "Measuring organizational legitimacy in social media: Assessing citizens judgments with sentiment analysis," *Business & Society*, vol. 57, no. 1, pp. 60–97, 2018.
- [18] S. Sohangir, D. Wang, A. Pomeranets, and T. M. Khoshgoftaar, "Big data: Deep learning for financial sentiment analysis," *Journal of Big Data*, vol. 5, no. 1, p. 3, 2018.
- [19] V. K. Neppalli, C. Caragea, A. Squicciarini, A. Tapia, and S. Stehle, "Sentiment analysis during hurricane sandy in emergency response," *International Journal of Disaster Risk Reduction*, vol. 21, pp. 213–222, 2017.
- [20] S. Goyal, "Review paper on sentiment analysis of twitter data using text mining and hybrid classification approach," *International Journal of Engineering Development and Research*, 2017.
- [21] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014.
- [22] D. Buscaldi and I. Hernandez-Farias, "Sentiment analysis on microblogs for natural disasters management: a study on the 2014 genoa floodings," in *Proceedings of the 24th International Conference on World Wide Web*. ACM, 2015, pp. 1185–1188.
- [23] G. Beigi, X. Hu, R. Maciejewski, and H. Liu, "An overview of sentiment analysis in social media and its applications in disaster relief," in *Sentiment Analysis and Ontology Engineering*. Springer, 2016, pp. 313–340.
- [24] Y. Ouyang, B. Guo, J. Zhang, Z. Yu, and X. Zhou, "Sentistory: multi-grained sentiment analysis and event summarization with crowdsourced social media data," *Personal and Ubiquitous Computing*, vol. 21, no. 1, pp. 97–111, 2017.
- [25] H. Shekhar and S. Setty, "Disaster analysis through tweets," in *Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on*. IEEE, 2015, pp. 1719–1723.
- [26] R. Srivastava and M. Bhatia, "Real-time unspecified major sub-events detection in the twitter data stream that cause the change in the sentiment score of the targeted event," *International Journal of Information Technology and Web Engineering (IJITWE)*, vol. 12, no. 4, pp. 1–21, 2017.
- [27] D. M. Blei, "Probabilistic topic models," *Communications of the ACM*, vol. 55, no. 4, pp. 77–84, 2012.
- [28] D. Barry and J. A. Hartigan, "A bayesian analysis for change point problems," *Journal of the American Statistical Association*, vol. 88, no. 421, pp. 309–319, 1993.
- [29] C. Erdman, J. W. Emerson *et al.*, "bcp: an r package for performing a bayesian analysis of change point problems," *Journal of statistical software*, 2007.