

Social media response and crisis communications in active shootings during COVID-19 pandemic

M. Ahmed Rusho^a, Md Ashraf Ahmed^b, Arif Mohaimin Sadri^{a,*}

^a Moss Department of Construction Management, Florida International University, Miami, FL 33174, USA

^b Department of Civil and Environmental Engineering, Florida International University, Miami, FL 33174, USA

ARTICLE INFO

Keywords:

Social media
Machine learning
Active shooting
COVID-19

ABSTRACT

Active shooting, a man-made hazard, remains an unsolved challenge as communities get threatened more frequently than ever before. As a low-cost alternative to the traditional approaches of responding to such crisis, data-driven approaches can help to identify more tailored response strategies and guide towards more informed decision-making. Recently, social media platforms helped researchers and practitioners with sufficient details and coverage to understand how communities respond to natural hazards through social media interactions. However, the empirical literature does not provide any comprehensive guidance on public reactions to active shootings as observed through social media interactions. This study adopted a holistic data analytics approach to collect large-scale social media data from Twitter (~252 K tweets, 04.17.20–05.20.20). The 2020 *Nova Scotia Attacks* were among the major shooting events observed during this period in addition to the unprecedented experiences people were having due to the COVID-19 pandemic. This study used several natural language processing and data mining approaches (such as temporal heatmaps, word bigrams, and topic mining) to cluster the social media crisis communication patterns of active shootings and create infographics of the diverse needs, concerns, and reactions observed in the aftermath of such events. Key interactions include bailing out of shooters, shooting investigation, police response, gun violence, lessons learned from the previous school (*Sandy Hook*) and mass shootings (*El Paso*), vehicle ramming (*Toronto Van Attack*), mobility issues, and health concerns during COVID-19 pandemic, changes in economy and education systems. This study would allow first responders and emergency management officials to enhance the capacity of social sharing platforms and facilitate risk communication in no-notice scenarios. Additionally, the infographics could serve as a data dictionary in future active shooting scenarios to maximize peer influence.

1. Introduction and motivation

Active shootings are man-made hazards that unfold quickly and can rarely be foreseen by the victims. The rising trend of active shooting incidents depicts it as one of the major threats to public safety. According to the Federal Bureau of Investigation, 779 fatalities and 1,418 injuries resulted due to active shootings in the USA between 2000 and 2017 (Engineering, 2019). In 2019, 417 active shootings were reported including high profile, mass casualty attacks, some of which occurred in less than 24 h interval. These shootings occurred in government territories, public spaces including movie theaters, nightclubs, shopping malls, religious facilities, hospitals that also depict the widespread occurrence of the phenomenon. Hence, the increasing tolls of casualties all over the world call for rapid solutions to this crisis.

In an active shooting scenario, the communication elements affiliated with the victims have profound implications in saving people's lives. Scholars agree that communication is the most important in crisis management (Mazer et al., 2015). The uncertainty and complex attributes of active shooting hazard necessitate the instigation of an effective communication tool for dissemination of the rapid onset of the event. Unfortunately, communication challenges persist in crisis moments due to the lack of preparedness and scholarly research focus in this arena. The widespread insurgence of active shooting incidents and their brutal consequences unfold the vulnerability and inefficiency of the existing emergency management plans.

Twitter is a microblogging platform that has introduced itself as a "real-time information network" among its users (Acar and Muraki, 2011; Hu et al., 2012). Twitter connectivity does not require user

* Corresponding author.

E-mail address: asadri@fiu.edu (A.M. Sadri).

<https://doi.org/10.1016/j.trip.2021.100420>

Received 8 October 2020; Received in revised form 15 June 2021; Accepted 18 June 2021

Available online 25 June 2021

2590-1982/© 2021 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

familiarity to share information. Thus it is prevalently used for news dissemination (Acar and Muraki, 2011). Expeditionous sourcing of mainstream media information (e.g. Osama Bin Laden killing) has gradually built credibility among its users (Hu et al., 2012). However, these communications take place informally among the users rather than the agencies. Currently, communication elements merely draw any attention to crisis emergency plans and are still being obliterated in the emergency management plans of the federal agencies. In recent years, people have been more engaged in social media platforms such as Twitter. Crisis onsets have been accentuated frequently over tweet discussions immediately after active shooting events. So it is high time these technologies are integrated into the crisis management plans. Consideration of social media platforms will ensure the early dissemination of shooting alerts as well as help the emergency management agencies to contribute to alleviating the tension surfacing on virtual communities.

As per Mak's study, the existing analyses have shortfalls in incorporating the long-term social media responses to extreme events and further do not apply diversified contextual approaches in the study (Mak and Song, 2019). Unlike other hazards, active shooting events have prolonged psychological impacts among people which needs to be addressed for comprehending social media data analysis. As a low-cost alternative to the traditional approaches of responding to such crises, data-driven approaches can help to identify more tailored response strategies and guide towards more informed decision-making. The empirical literature does not provide any comprehensive guidance on public reactions to active shootings as observed through social media interactions. As such, this study utilizes a holistic approach to understand the social media response and crisis communications in active shootings with pandemic (COVID-19) concurrence.

This study adopted a holistic data analytics approach to collect large-scale social media data from Twitter (~252 K tweets, 04.17.20–05.20.20). The 2020 *Nova Scotia Attacks* were among the major shooting events observed during this period in addition to the unprecedented experience people were having due to the COVID-19 pandemic. This study used several natural language processing and data mining approaches (such as temporal heatmaps, word bigrams, and topic mining) to cluster the social media crisis communication patterns and create infographics of the diverse needs, concerns, and reactions observed in the aftermath of such events. The infographics presented in this study could serve as a data dictionary in future active shooting scenarios to maximize peer influence.

2. Background and related work

Media dependency theory is a concept that predicts the need for mediated information during any crisis moment (Mazer et al., 2015). The theory is reliant on user's trust profiling criteria. The topology of the social media network and the behavior of the social media users support the principles of the theory. However, situational factors varying across incidents influence these need patterns. During catastrophic moments, crisis communication aims at reducing the negative impacts of the havoc (Herovic et al., 2020). Communication media ensures the convergence of the messages and coordination of the resources. Media dependency theory features advocate for the facts and enforce reliance on preferred sources with more weightage in comparison to mediated sources in crisis moments.

Since collaboration and communication in social media are more effective during crises, social media users exploit the opportunity rapidly. Social media dependency as a communication tool during an active shooting scenario can be apprehended in two aspects. Firstly, social media platforms are widely used, and one click away for access around the world. Secondly, social media is a readily available tool for alarm dissemination without incurring any time for source authentication (Bunker et al., 2017). Nowadays, people are more connected to social media rather than mainstream news channels. Platforms like

Twitter, Facebook, etc. work as the global hub for instant response and spread of news. Sharing information in social media assures connectivity among the users with better awareness and higher visibility (Acar and Muraki, 2011). During an active shooting phenomenon, flickering real-time information is crucial in avoiding the exacerbation of the unrest situation. Hence, social media can be used as an effective alarm dissemination tool in emergency communication management.

The recent studies also speak for the prevalence of social media activity. Platforms like Facebook, Twitter, etc. act as a resource for mass opinion formation and emotional guidance during crisis moments (Liu et al., 2011). As per recent statistics, Twitter users generate 143 K tweets per second with the containment of trending topics (Sadri et al., 2017a). In previous active shooting cases, significant collective communication patterns have been observed, e.g. shooting at the University of Texas (Li et al., 2011). As per the study from Lin Tzy Li, the followers of UT-Austin on Twitter had an outburst of emotions immediately after the shooting happened. The study also showed the transition of discussion trends throughout the timeline. There are instances when Twitter has been more effective in dispersing information of national interest more rapidly than conventional news sources (e.g. Osama Bin Laden killing) (Hu et al., 2012). So, social media platforms have the potential to reshape the prospects of existing communication tools. Unlike the local news context, social media has observed the outcry due to the active shooting rampage beyond borders. In the future, social media analytics will help us to provide deeper insights into these behavioral sets to configure strategies that are malleable to multiple needs.

Misinformation spreading and detection in social media are also becoming an emerging concern. Battur et al., detected Twitter bots, which is a software that sends fake tweets automatically to users. Detecting bots is necessary to identify fake users and to protect genuine users from misinformation and malicious intentions. Asr et al., stated that misinformation detection at the level of full news articles is a text classification problem and reliably labeled data in this domain is rare. Previous work relied on news headlines, microblogs, tweets, and articles collected from so-called "reputable" and "suspicious" websites and labeled accordingly. Authors leveraged fact-checking websites to collect individually labeled news articles concerning the veracity of their content (Asr and Taboada, 2018).

Besides, multiple communities located in proximity to the shooting locations may directly or indirectly be impacted due to such no-notice events. Most often, first responders need to secure the perimeter of the shooting premises intervening in regular traffic flow. Such communities, with too many uncertainties, need to respond to the crisis event with effective information dissemination as part of management, planning, and promoting situational awareness of residents to experience reduced threats (Sadri et al., 2018). In major active shooting events, evacuation of nearby communities may be required, or traveling to the affected areas may be restricted. For example, in the 2020 *Nova Scotia Shootings*, the shooter traveled to multiple places to kill people randomly. So, the lawmaker had to impose a sudden restriction abrupting the regular traffic network. In cases like this, individuals' relationship with social partners and social network characteristics define their safety and determine: (a) how the risk is perceived, (b) how fast the information is received (Sadri et al., 2017b). So, the diffusion of situational awareness can play a vital role in active shootings to control the underlying networks of followees and followers that exist in social media platforms (Roy et al., 2020b). Broadcast of salient Twitter features like sheltering, evacuation, hazard locations, counseling, etc. can help in leveraging the risk communication structure and contribute towards community resilience (Ukkusuri et al., 2014).

3. Data description

In this study, the analyses were conducted on the clean tweets collected through Twitter Streaming API. On Mar 27, 2020, the data collection was conducted on a trial basis. Based on the satisfactory

results derived from the trial data, the data collection process was run for around a month (Apr 17, 2020, to May 20, 2020). Both of these data sources have been merged in the analyses. As the study focuses on the apprehension of people's responses and reactions over the active shooting incidents all over the world, keyword-driven data collection was preferred to location-specified data collection. Potential keywords were listed initially after the authors' discussion. Tweets were searched manually with each potential keywords for observing their relevance to the topic. Finally, keywords producing irrelevant tweets were discarded from the list. The final set of keywords consisted of *active shootings*, *activeshooting*, *mass shootings*, *massshooting*, *school shootings*, *schoolshooting*, *gun violence*, *gunviolence*, *shootings*, *shooter*. The plural form of the keywords (e.g. *active shootings*, *mass shootings*) generated more relevant tweets than the singular forms. For data comprehension, both of the forms were considered. Keywords like *shooting*, *shoot*, *gun policy* were discarded as they fetched irrelevant topics.

The raw dataset included ~ 252,000 tweets for the aforementioned timeline. The percentage of location-enabled tweets was pretty low. Since the study does not focus on the spatial aspects of active shootings it did not impact the analyses. The follower counts of the tweet IDs ranged up to ~ 5.7 million that indicates the participation of highly influential Twitter accounts in the active shooting discussions. The raw tweet database is comprised of both English and Spanish tweets. Spanish tweets were trimmed off the dataset. The clean dataset was also filtered from stopwords (e.g. hashtags, weblinks, hallmarks, tweet handles, etc.). The clean dataset included ~ 157,000 tweets which were considered as the core data. The subsequent datasets were filtered from bad words (e.g. words with spelling mistakes, no meaning, or offensive meaning) to omit erroneous results.

4. Methodology

The section comprises a brief explanation of methods used for developing the data analytics and infographics in this study.

4.1. Word frequency and temporal heatmaps

Word frequency is a prevalently used natural language processing technique that counts the appearances of words for a particular text assembly and provides a sequential ordering of the counts. The temporal heatmap is an advanced pivoting technique used to demonstrate the word frequency over a specified time-series. The frequency is represented with a color bar chart. The deviations in color sequencing help to differentiate the frequency of the words. In this study, the top "100" most frequent words have been used to generate the temporal heatmap with alphabetical sorting.

4.2. Word N-grams analysis

Word n-grams is a natural language processing technique used to define the word probability of "n" number of words appearing in a word sequence. It trims word pairs or triplets (i.e. for bigrams and trigrams) as a single entity and identifies next word probability in the text assembly (Kumar, 2017).

The word sequence $w1...wN$ is separated from the order of text observations $x1...xT$ for which the posterior probability $Pr(w1...wN | x1...xT)$ achieves the maximum value. It can be written in the following form :

$$\arg \max \{ Pr(w1...wN | x1...xT), Pr(x1...xT | w1...wN) \}, w1...wN \quad (1)$$

where, $Pr(x1...xT | w1...wN)$ is the conditional probability of, the given word sequence $w1...wN$. We obtain the decomposition using the following conditional probabilities:

$$Pr(w1...wN) = wn | w1 \dots wn-1 \quad (2)$$

We can partition the vocabulary of size W in number G of word classes. The category mapping can be written as:

$$G: w \rightarrow Gw \quad (3)$$

Here each word w of the vocabulary is mapped to its word-class Gw . For a word bigram (v, w) we use (Gv, Gw) for denoting the corresponding class bigram. For maximum likelihood estimation, the equation could be written as:

$$Fbi(\beta) = N(w) + \quad (4)$$

Given,

Fbi = bigram maximum likelihood

W = vocabulary size

u, v, w, x = words in a running text.

N = training corpus size

β = number of word classes (Martin et al., 1998)

4.3. Topic modeling

The topic model is a mathematical model derived from machine learning and natural language processing that provides the concepts of topics appearing in a set of text information. It is a popular text-mining tool that explores the intrinsic framework of a text body. A large volume of unsupervised text-data can be analyzed with topic modeling (Blei et al., 2010). Blei et al. (2003) developed Dirichlet priors for model generation and learned θ and ϕ by maximization of the probability of collection words by using variational Bayes. Griffiths and Steyvers (2004) proposed the use of collapsed Gibbs sampling for making an inference approximation by the posterior distribution over the assignments of words to topics ($P(z|w)$). To summarize the iteration, the latest function for a new assignment topic in the process of sampling is expressed as follows:

$$P(z = t | z, w, \alpha, \beta) \propto \frac{n(d, t) + \alpha}{n(d, t) + T\alpha} \frac{n(t, w) + \beta}{n(t, w) + W\beta} \quad (5)$$

where, $n(d, t)$ is the number of assignments of topic t in document d , and $n(t, w)$ is the number of assignments of word w to topic t ; all counts exclude the current assignment z . The variables' descriptions are listed as follows:

T = Number of topics W = Number of unique words (vocabulary) D = Number of documents N = Number of tokens $\theta = D \times T$ of probabilities; Topic distribution in documents $\phi = T \times W$ of probabilities; Word distribution in topics $\alpha = D \times T$ of α priors; Dirichlet prior for $\theta \beta = T \times W$ of β priors; Dirichlet prior for $\phi w = N$ -Vector of word identity w ; Words in documents $z = N$ -Vector of topic assignment z ; Topic Assignment of Words (Lau et al., 2012).

5. Discussion of findings

Different machine learning techniques were used to define the diversified characteristics of the data source. Temporal heatmaps with frequent words, word network analysis with n-grams, topic modeling are the techniques that were used to model the outputs. Each of these techniques provided unique insights and added different dimensions to the results. A synopsis of these methods is discussed in the following sections:

5.1. Temporal heatmaps with frequent words

Frequent word analysis is a conventional approach of sorting the words that have appeared the most number of times in a pool of words. Words that have buzzed higher certainly provide the gists of discussion on social media platforms. Heatmapping those frequent words allow us to visualize how these words are stranded in a given period. It helps to identify how these words surge and perish within the timeframe. In this

study, the top 100 most frequent words have been computed as listed in Table 1.

Temporal heatmaps have been plotted for these top 100 most frequent words and illustrated into two figures (Fig. 1 and Fig. 2). The words have been tallied in alphabetical order. These words can serve as an active shooting “dictionary” as it incorporates all the aspects of discussion that can emerge in discussions. The word frequency scale varies from the most frequent word *shooter* (~48 k appearances) to the 100th frequent word *crime* (~1.8 k appearances). The top few rows of the most frequent words show some obvious findings with the keywords. However, the pattern in which they appear in the heatmap gives an exciting impression of the data. These words are crowded around two timeslots.

The first timeslot is between April 18, 2020, to April 19, 2020. It has the exact timeline of the *Nova Scotia* shooting (Collins, 2020). Frequent words like *gabriel*, *wortman*, *suspect* depict people’s tension during that time. Co-occurrence of other words like *killed*, *rcmp*, *investigation* also expresses people’s concern over the role of Royal Canadian Mount Police (RCMP) during the event. However, the quick perishment of these words also indicates that people’s reactions to the active shooting events are

Table 1
Ranking and Frequency of 100 Most Frequent Words.

Rank	Words	Count	Rank	Words	Count
1	shooter	48,218	51	want	3,475
2	gun	24,648	52	kent	3,375
3	violence	21,297	53	pensacola	3,273
4	shootings	20,307	54	end	3,229
5	active	16,515	55	yet	3,184
6	response	14,629	56	old	3,136
7	school	14,283	57	dylan	3,098
8	like	12,425	58	th	2,924
9	elementary	12,091	59	white	2,906
10	people	9,742	60	times	2,873
11	children	9,618	61	guns	2,869
12	drills	8,993	62	world	2,792
13	trauma	8,913	63	need	2,760
14	expect	8,824	64	still	2,722
15	endure	8,759	65	custody	2,599
16	freedoms	8,665	66	son	2,580
17	year	8,598	67	america	2,537
18	much	8,520	68	get	2,534
19	prevent	8,319	69	anniversary	2,528
20	others	7,719	70	classroom	2,526
21	mass	7,594	71	know	2,488
22	quote	7,593	72	black	2,461
23	easy	7,464	73	information	2,431
24	virus	7,450	74	released	2,418
25	seems	7,359	75	students	2,412
26	killed	7,299	76	us	2,392
27	feels	7,209	77	time	2,388
28	shot	7,188	78	would	2,383
29	police	6,056	79	power	2,376
30	years	5,369	80	lead	2,366
31	better	5,334	81	lebron	2,366
32	one	5,096	82	say	2,355
33	future	5,059	83	new	2,295
34	scotia	4,783	84	create	2,274
35	nova	4,742	85	said	2,267
36	suspect	4,597	86	american	2,201
37	wortman	4,447	87	marks	2,194
38	shooting	4,442	88	right	2,088
39	investigation	4,407	89	day	2,023
40	amp	4,388	90	go	1,998
41	state	4,366	91	injured	1,993
42	gabriel	4,358	92	free	1,982
43	today	4,267	93	four	1,967
44	rates	4,030	94	may	1,940
45	sandy	3,738	95	situation	1,924
46	hook	3,728	96	many	1,916
47	died	3,725	97	killing	1,864
48	rcmp	3,721	98	stop	1,852
49	least	3,718	99	university	1,850
50	fbi	3,604	100	crime	1,848

rather ephemeral.

The second timeslot of tweet word crowding occurred throughout the first week of May 2020. In the 1st week of May, multiple shooting events took place. Even though none of these events created mass casualties the panic surfacing throughout social media was evident with the high number of tweets. The words in the tweets were mostly related to high school shootings. Surprisingly, this revealed that people reminisce about historical shooting events to anticipate their degree of distress to the current events. For instance, *Kent State Shootings* occurred in 1970 (Caputo, 2005), and *Sandy Hook Elementary School Shootings* occurred in 2012 (Scinto, 2012) but they both appeared in discussions about modern-day shootings. Nevertheless, a few words like *American*, *America*, *US* grabbed people’s attention constantly. This also means that the United States is also the “talk of the table” due to its high active shooting rates. On the other hand, words like *pensacola* denoting the “Naval Air Station Pensacola Shooting” of 2019 (Carrega et al., 2019) gained sudden momentum at the end phase of the data collection.

5.2. Word network mapping and N-grams

N-gram is a technique used in machine learning to compute the sequence of words where N is an integer value. Hence, N-gram is a sequence of N number of words (Kumar, 2017).

For example, if the words *active shooting incidents* appear in the text database N-grams can be computed for up to N = 3 words. For, N = 2 two N-grams can be formed: *active shooting* and *shooting incidents*. Whereas, for N = 3 only one N-grams can be formed which is the sequence of words itself. The N-grams are also called “bigrams” and “trigrams” respectively for N = 2,3 values. The most common bigrams and trigrams in a text database can help in identifying the most vocal topics of social interactions. In this study, the top 100 bigrams and trigrams have been computed for analyses which are listed in (Tables 2 and 3).

In many instances, it becomes difficult to track the right set of topics with only frequent word analysis. As bigrams and trigrams are word pairs and triplets co-appearing in a pool of words, they provide a logical explanation of those word frequencies. For example, frequent words like *th*, *four*, or *dylan* do not provide a complete scenario of the associated topics. With bigrams like (*th*, *anniversary*) it is clear that the bigram is related to the anniversary of the shooting events. Similarly, trigrams like (*four*, *students*, *killed*) define that the frequent word *four* is directed to the school shootings. Tweets like “*Today marks 50 years since the shootings at Kent State on May 4, 1970. To honor this significant moment in American history*” reinforce the fact that n-grams help in providing substances of the recurrent discussions. Sometimes, a comparison of bigrams and trigrams for similar words is required for a precise elaboration of bigrams. For example, bigrams like (*apple*, *declined*), or (*small*, *silver*) are difficult to explain. If similar trigrams are analyzed it shows that the (*apple*, *declined*) bigram is related to the Pensacola shooting event where the FBI succeeded to decrypt the communication device (I phone) of the shooter (Carrega et al., 2019). Similarly, (*small*, *silver*) is related to the suspect of the Nova Scotia shooting who was driving a silver car during the shooting event (Collins, 2020). Hence, surfing through the trigrams with similar words provides an explicit answer to these queries.

Word bigrams can be used to develop a map of word networks where the words work as nodes and bigrams work as connectors. This is an efficient way to explore the inter-connectivity of the bigrams and how different bigrams stem through the network.

In a network, nodes that have central positions have more connections compared to the edge nodes. Likewise, in a word network, words that are relatively in central positions have more connectivity with other words. This also depicts how different topics have branched out from the same word node and propagated throughout the network. After multiple iterations for best visualization, the word network map has been plotted for the top 50 bigrams in this study. (Fig. 3) shows two major and a few minor word networks. *School* acts as a central or “root” node for three

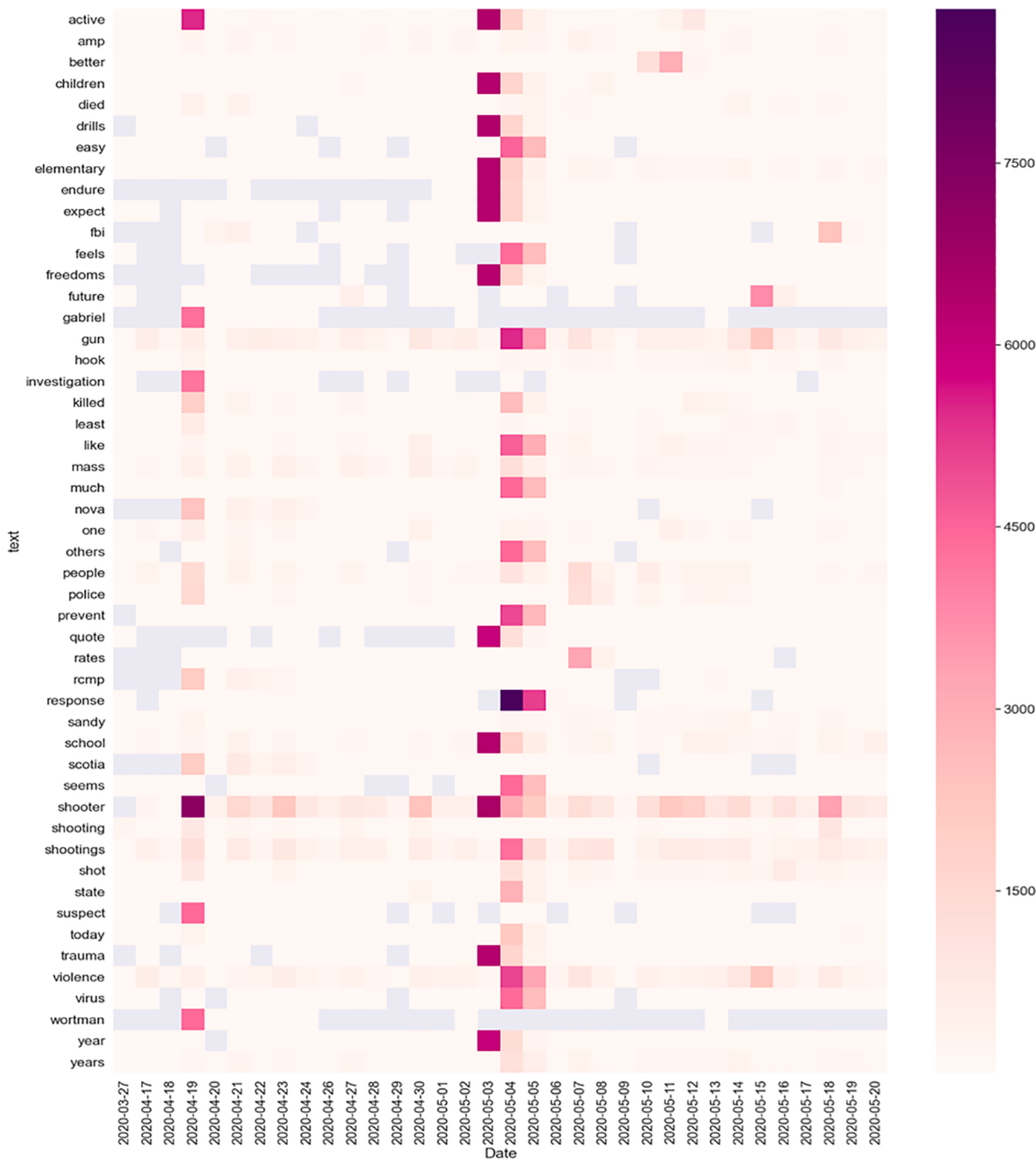


Fig. 1. Temporal heatmaps for top 1 to 50 words.

school shooting topics. *Died* also acts as a central node which is not only affiliated to different active shooting incidents but also with discussions on the psychological impact of active shootings on common people. The centrality and multiple connectivities of words like *school* and *died* indicate their impact on users' perceptions and concerns. *Gun* is the central node for the other word network which stems through two different topics: (1) people's perception of gun violence, (2) people's response to gun topics during the pandemic of coronavirus (mentioned as *virus*).Fig. 4.

5.3. Topic modeling

Topic modeling is a statistical model that generates abstracts of topics using a probabilistic process based on a set of sampling rules

(Sadri et al., 2018). This process explains how topics are generated in a text database based on some latent variables. The model performance can be diagnosed based on certain parameters like perplexity, log-likelihood, coherence value, and so on. In this study, the authors evaluated the coherence value for topic model optimization. Each topic model has an overall coherence value including an individual score for each topic within the model. The higher the coherence value the better will be the performance of the model. After the coherence value touches the peak, the topics start recurring in the model affecting its efficiency. In this study, the topic model has been diagnosed with the coherence value for 1 to 75 topics. The highest score of 0.578421 has been obtained for 47 topics.

For each topic, 30 salient words were generated from which 10 significant words have been reported in Table 4. As the social media

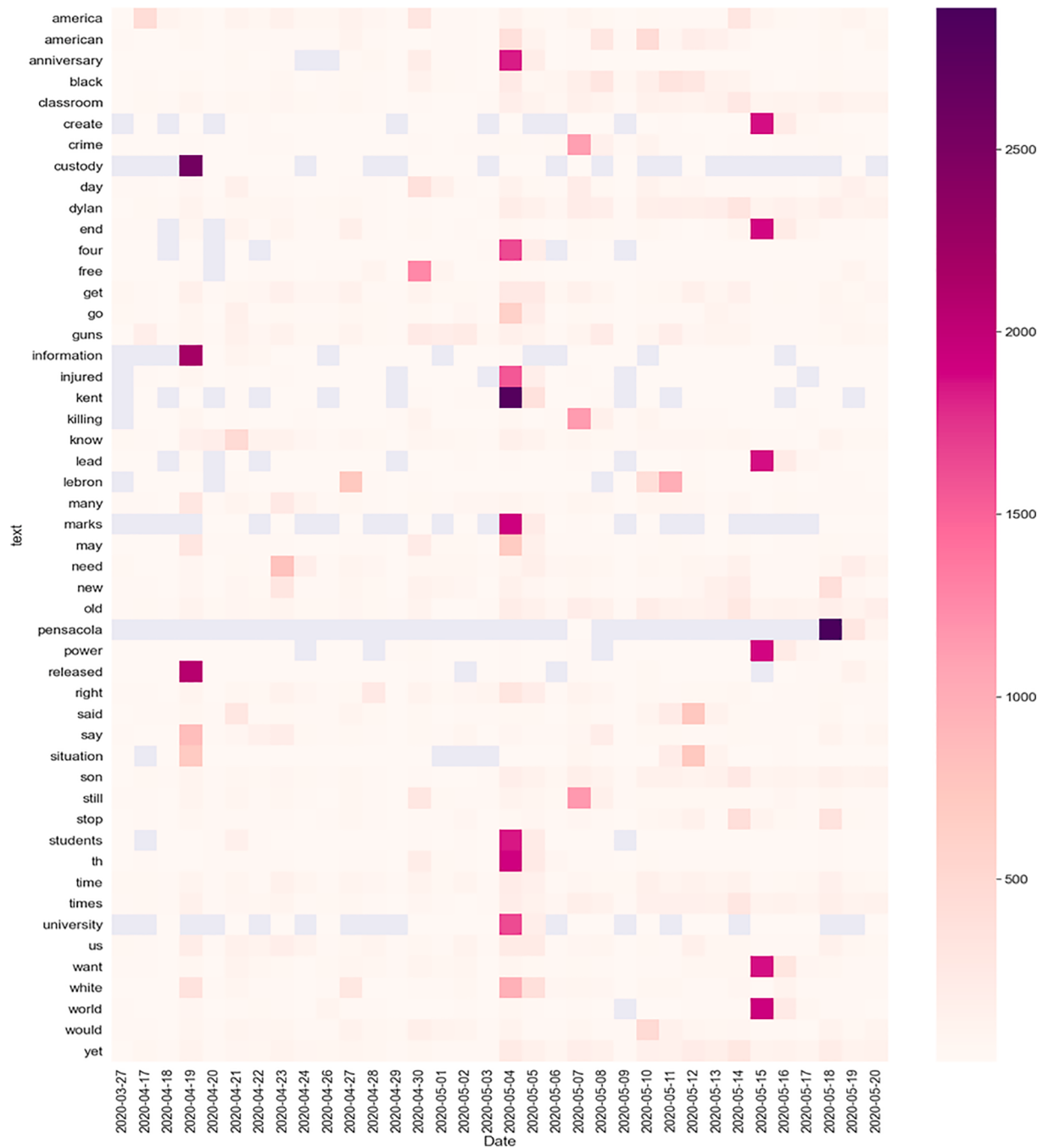


Fig. 2. Temporal heatmaps for top 51 to 100 words.

response was collected during the crisis of COVID-19, several topics related to COVID-19 were observed. Surprisingly, some of the active shooting incidents and facts were unveiled with topic modeling that were not captured with the prior analysis techniques. The relevant topics can be channelized into three mainstreams: (1) topics related to the active shooting facts and incidents (2) topics related to the COVID-19 (3) topics related to the effects of COVID-19 on active shootings. Highlights of the prominent topics are discussed below:

5.4. Dynamic topic model (Over Time)

The dynamic topic model is a probabilistic time-series model that shows the evolution of the topic and word distribution in the topic over any specific timestamp (Ahmed et al., 2020; Roy et al., 2020a). In Fig. 1,

most of the active shooting-related words appeared from May 3 to May 5 on Twitter. Hence, the focus of this section is to explore the topic dynamics over this timeline. In these three days, more than 36,000 tweets have been found on which dynamic topic model is performed. From the coherence values of topics of this sub-dataset, 11 optimum number of topics have identified. Fig. 5 is representing the most frequent words in each topic over these three days timestamp.

People started discussing gun violence (topic 1, topic 4) and shooting (topic 5, topic 6) on May 3 as situational awareness of a recent shooting event happened just two weeks after (Nova Scotia shooting). The impact of the coronavirus pandemic along with gun violence was also discussed (topic 1) on May 3. In next day (May 4), one shooting event was mentioned (topic 0) where a white racist teenager shot and killed someone's daughter. In topic 3, people were emphasizing "liberal

Table 2
Ranking and Frequency of 100 Most Frequent Bigrams.

Rank	Bigrams	Count	Rank	Bigrams	Count
1	(gun, violence)	17,961	51	(want, future)	2,151
2	(active, shooter)	16,381	52	(future, end)	2,149
3	(shooter, drills)	8,895	53	(violence, lead)	2,145
4	(elementary, school)	8,705	54	(today, marks)	2,136
5	(endure, trauma)	8,701	55	(custody, information)	2,086
6	(trauma, active)	8,700	56	(information, released)	2,066
7	(school, children)	8,699	57	(state, university)	1,815
8	(children, endure)	8,694	58	(shot, killed)	1,760
9	(expect, elementary)	8,690	59	(anniversary, shootings)	1,740
10	(drills, freedoms)	8,612	60	(students, killed)	1,740
11	(year, expect)	7,555	61	(marks, th)	1,727
12	(quote, year)	7,554	62	(four, students)	1,718
13	(response, gun)	7,153	63	(nine, injured)	1,684
14	(violence, seems)	7,124	64	(killed, nine)	1,677
15	(much, like)	7,121	65	(university, four)	1,676
16	(like, response)	7,111	66	(shooter, situation)	1,629
17	(virus, response)	7,109	67	(better, shooter)	1,580
18	(seems, easy)	7,108	68	(shooter, better)	1,480
19	(easy, prevent)	7,107	69	(better, passer)	1,440
20	(response, feels)	7,106	70	(better, rebounder)	1,429
21	(feels, much)	7,106	71	(passer, better)	1,427
22	(prevent, others)	7,105	72	(top, head)	1,423
23	(mass, shootings)	4,564	73	(faster, better)	1,422
24	(gabriel, wortman)	4,345	74	(lebron, stronger)	1,421
25	(shooter, investigation)	4,187	75	(stronger, taller)	1,421
26	(nova, scotia)	4,149	76	(taller, faster)	1,421
27	(suspect, active)	4,085	77	(head, lebron)	1,420
28	(wortman, suspect)	3,860	78	(sharp, shooter)	1,412
29	(sandy, hook)	3,708	79	(free, throw)	1,379
30	(hook, elementary)	3,363	80	(killing, people)	1,365
31	(kent, state)	3,293	81	(throw, shooter)	1,359
32	(school, shootings)	3,211	82	(student, section)	1,323
33	(end, gun)	2,567	83	(pensacola, shooter)	1,323
34	(years, old)	2,564	84	(section, distracted)	1,317
35	(son, dylan)	2,521	85	(distracted, free)	1,317
36	(least, times)	2,521	86	(shot, shooter)	1,313
37	(times, classroom)	2,521	87	(crime, rates)	1,296
38	(dylan, shot)	2,520	88	(police, still)	1,283
39	(shot, least)	2,520	89	(people, killed)	1,280
40	(classroom, sandy)	2,520	90	(still, killing)	1,275
41	(elementary, died)	2,520	91	(arrest, rates)	1,272
42	(died, years)	2,520	92	(rates, arrest)	1,271
43	(old, yet)	2,520	93	(rates, incarceration)	1,271
44	(th, anniversary)	2,379	94	(incarceration, rates)	1,271
45	(investigation, custody)	2,223	95	(rates, police)	1,271
46	(lead, world)	2,156	96	(saudi, shooter)	1,040
47	(create, future)	2,153	97	(run, active)	1,029
48	(shootings, kent)	2,152	98	(said, run)	1,028
49	(power, create)	2,152	99	(apple, declined)	1,021
50	(future, want)	2,152	100	(small, silver)	1,009

snowflakes” who were trying to prevent active shootings by spreading awareness on May 4. Snowflake is a political insult for someone who is perceived as too sensitive, often used for millennials and liberals (Dictionary, 2020).

Besides, a common quote is found (topic 9) about the trauma experienced by elementary school children due to active shooting drills on the same day, which was retweeted thousands of times. Also, a previous shooting event at Kent State (topic 5), where four students were killed got notified on May 4. In next day (May 5), people mentioned about the white shooter and the influence of COVID19 on active shooting (topic 10). In topic 4, an Illinois carpenter, Greg Zanis was remembered who built 27,000 handmade white crosses to honor the victims of US active shooting events on the same day. Lastly, racism in shooting events (black victims) was discussed on May 5. So, the summary is people were concerned about the gun violence, previous shooting events, victims, racism

Table 3
Ranking and Frequency of 100 Most Frequent Trigrams.

Rank	Trigrams	Count	Rank	Trigrams	Count
1	(active, shooter, drills)	8,836	51	(custody, information, released)	2,064
2	(endure, trauma, active)	8,699	52	(kent, state, university)	1,801
3	(trauma, active, shooter)	8,698	53	(th, anniversary, shootings)	1,739
4	(school, children, endure)	8,694	54	(marks, th, anniversary)	1,726
5	(children, endure, trauma)	8,692	55	(anniversary, shootings, kent)	1,701
6	(elementary, school, children)	8,690	56	(four, students, killed)	1,692
7	(expect, elementary, school)	8,690	57	(today, marks, th)	1,687
8	(shooter, drills, freedoms)	8,612	58	(students, killed, nine)	1,677
9	(year, expect, elementary)	7,556	59	(state, university, four)	1,676
10	(quote, year, expect)	7,547	60	(university, four, students)	1,676
11	(response, gun, violence)	7,155	61	(killed, nine, injured)	1,667
12	(like, response, gun)	7,107	62	(active, shooter, situation)	1,611
13	(seems, easy, prevent)	7,107	63	(better, shooter, better)	1,436
14	(gun, violence, seems)	7,107	64	(better, passer, better)	1,427
15	(feels, much, like)	7,106	65	(shooter, better, rebounder)	1,422
16	(virus, response, feels)	7,106	66	(faster, better, passer)	1,422
17	(response, feels, much)	7,106	67	(passer, better, shooter)	1,422
18	(much, like, response)	7,106	68	(stronger, taller, faster)	1,421
19	(easy, prevent, others)	7,105	69	(top, head, lebron)	1,420
20	(violence, seems, easy)	7,105	70	(head, lebron, stronger)	1,420
21	(drills, freedoms, quote)	4,968	71	(lebron, stronger, taller)	1,420
22	(freedoms, quote, year)	4,967	72	(taller, faster, better)	1,420
23	(active, shooter, investigation)	4,174	73	(free, throw, shooter)	1,358
24	(suspect, active, shooter)	4,085	74	(student, section, distracted)	1,317
25	(wortman, suspect, active)	3,842	75	(distracted, free, throw)	1,317
26	(gabriel, wortman, suspect)	3,802	76	(section, distracted, free)	1,317
27	(sandy, hook, elementary)	3,362	77	(still, killing, people)	1,274
28	(shot, least, times)	2,520	78	(incarceration, rates, police)	1,271
29	(dylan, shot, least)	2,520	79	(rates, police, still)	1,271
30	(died, years, old)	2,520	80	(rates, incarceration, rates)	1,271
31	(times, classroom, sandy)	2,520	81	(crime, rates, arrest)	1,271
32	(son, dylan, shot)	2,520	82	(police, still, killing)	1,271
33	(classroom, sandy, hook)	2,520	83	(rates, arrest, rates)	1,271
34	(years, old, yet)	2,520	84	(arrest, rates, incarceration)	1,271
35	(least, times, classroom)	2,520	85	(lead, world, power)	1,260
36	(elementary, died, years)	2,520	86	(world, power, create)	1,254
37	(hook, elementary, died)	2,520	87	(information, released, gabriel)	1,224
38	(end, gun, violence)	2,494	88		1,224

(continued on next page)

Table 3 (continued)

Rank	Trigrams	Count	Rank	Trigrams	Count
39	(shooter, investigation, custody)	2,209	89	(released, gabriel, wortman)	1,029
40	(power, create, future)	2,152	90	(said, run, active)	1,028
41	(create, future, want)	2,152	91	(shooter, investigation, believed)	1,006
42	(future, want, future)	2,149	92	(investigation, believed, driving)	992
43	(future, end, gun)	2,148	93	(believed, driving, small)	991
44	(want, future, end)	2,148	94	(driving, small, silver)	991
45	(gun, violence, lead)	2,146	95	(succeeded, unlocking, encryption)	977
46	(violence, lead, world)	2,143	96	(fbi, succeeded, unlocking)	974
47	(shootings, kent, state)	2,139	97	(shooter, iphone, apple)	974
48	(investigation, custody, information)	2,086	98	(unlocking, encryption, saudi)	973
49	(others, virus, response)	2,082	99	(encryption, saudi, shooter)	973
50	(prevent, others, virus)	2,080	100	(pensacola, fbi, succeeded)	973

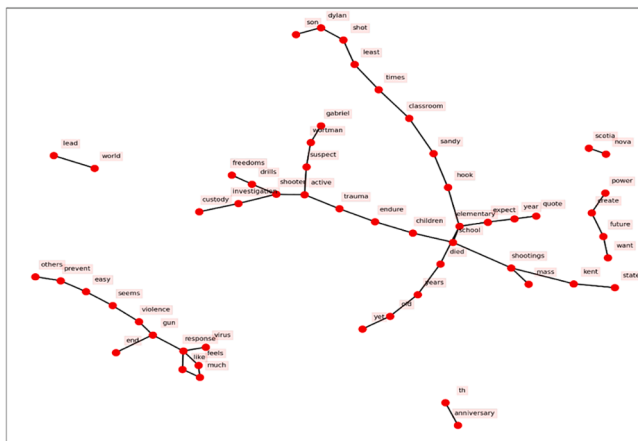


Fig. 3. Work network map for the top 50 bigrams.

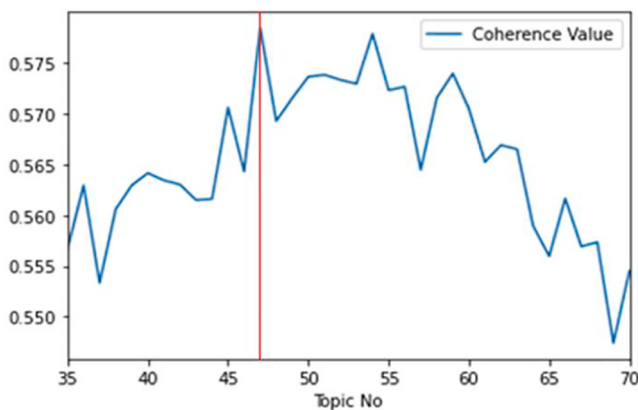


Fig. 4. Coherence value vs topic no.

Table 4

Highlights of Topic Modeling.

Topic Stream	Probable Topic	Ten Salient Words in Topic	Brief Explanation
Active shooting facts and incidents	Recent Bails of Shooters	shooter, response, released, amid, global, legit, superior, vulnerable, disarming, weaponry	The changes in incarceration policies for the shooters that raise the vulnerability of the public safety
	Gun Violence History	gun, violence, contact, stronger, history, victim, sense, launch, threat, business	The existence of gun violence in the past and their effects and threats on the economy (Silverstein, 2020 ; GVA, 2021)
	The Insurgence of School Shootings	school, rate, know, forget, still, crime, become, high, imagine, age	How school shootings are becoming common day by day including juvenile shooters and victims
	Nighttime Crimes	power, night, neighborhood, expect, easy, key, murder, thought, officer, friend	The higher crime rates, murders, shootings that take place during the late hours and the likelihood of their occurrence
	Changes in the Shooting Investigation Process	active, investigation, change, despite, listening, biggest, movement, charged, zealand, weaponry	Jurisdictional leniency towards the mass shooters. <i>Zealand</i> is mentioned that indicates to the “Christchurch Mosque Shooting” event
	El Paso Shooting	month, last, white, call, safety, big, remember, threat, blood, el, paso	The brutality of the 2019 El Paso shooting at Walmart, Texas which killed 23 people and injured 23 others as a part of white nationalist and anti-immigrant theme (Eligon 2019)
	Houston Police Shooting	died, life, mom, lost, saw, crisis, houston, speak, currently, survive	Killing of an arrestee by the police in front of his family on April, 21,2020 while he was surrendered on his knees. The news created grief among the public (Blakinger, 2020).
The effects of COVID-19 on active shootings	Toronto Van Attack	terrorist, actual, daniel, example, justice, toronto, boy, citizen, reform, identify	Killing spree of Alek Minassian, a 25-year-old man from Ontario with a van at Toronto, on April 23, 2018 (Tait, 2018)
	Shootings During Lockdown	people, shooting, mass, lockdown, arrest, happen, miss, true, learning, ask	The upsurge in shooting arrests during the lockdown and the learnings from these events.
	Police Response During Pandemic	police, seems, group. join, emergency, watch, pandemic, fire, cop, alert	The extra vigilance and support from law enforcement officers during emergencies in this pandemic situation
	Gun Violence in the USA During COVID-19	one, trump, Chicago, coronavirus, supporter, portland,	Abrupt spike of shooting events in Portland; 22

(continued on next page)

Table 4 (continued)

Topic Stream	Probable Topic	Ten Salient Words in Topic	Brief Explanation
COVID-19		gunviolence, receive, trust, yesterday	shootings between March 29 to April 9, 2020 (Levinson, 2020)
	Changes in Economy and Education System Due to COVID-19	lead, today, run, ban, virus, pm, student, order, online, worry	Changes and impacts in different sectors and adaption of online modules due to coronavirus
	Public Movements During Lockdown	first, day, every, public, important, march, april, texas, break, rule	How social distancing guidelines are violated in different states, i.e. Texas
	Health Security Concerns at Workplace During Coronavirus	work, hit, must, sure, health, covering, essential, stopped, mental, physical	The necessity of protective equipment to fight coronavirus and the mental and physical impacts of lockdown
	Community-Based Protective Measures for COVID-19	protect, community, death, covid, weapon, neighbor, member, town, sport, panic	Fighting coronavirus from a community perspective and taking protective measures accordingly

in shooting, spreading awareness about the active shooting, and the influence of the coronavirus pandemic when any shooting event occurs.

6. Conclusions

Active shooting is one of the emerging crises that remains an enigma to policymakers and first responders. The unpredictable nature and severity of the phenomenon necessitate contemporary measures for such no-notice man-made hazards. The existing literature does not provide sufficient guidance towards enhancing community resilience observed through social media-based interactions of risk and response in the aftermath of active shootings. In crisis moments like active shootings, influence through such social sharing platforms tends to escalate and communication of risks may guide or mislead people in taking certain actions. Machine learning techniques and data-driven approaches can help to analyze the intrinsic motives, thoughts, and sentiments retrieved through social media interactions and facilitate more informed decision-making as part of developing more efficient response strategies.

As such, this study adopted data-driven approaches and accumulated tweets over four consecutive weeks during the peak of the COVID-19 outbreak with the same set of keywords related to active shootings. During this period we also observed catastrophic shooting events such as the 2020 *Nova Scotia Attacks*. In our analyses, we exploited the active shooting related virtual discussions on Twitter from multiple natural language programming aspects: (i) 100 most frequent words capture the most buzzed words in the discussions, (ii) temporal heatmaps capture how the topic emerge and trend through the timeline during and after the shootings, (iii) word network diagrams and 100 most frequent bigrams and trigrams capture the centrality and interconnectivity of the topics, (iv) topic mining shows the COVID-19 pandemic induced active shooting discussions and how they influence the scenario (i.e. police vigilance, increased rate of the prisoner release, etc.), (v) dynamic topic mining illustrates how the topics perish or intensify in a specified timeframe. For each of these methods, different units of analysis have been utilized and the units to exhibit the best possible outcome have been reported. For example, the top 100 bi-grams and tri-grams have been reported in the study. Further expansion of the reporting units fetches illogical datasets which contrast the scope of the analysis. Similarly, the optimum unit of analysis for the topic modeling has been

derived by the use of maximum coherence value. These methods helped to cluster public responses and interactions following major active shooting scenarios. Such key insights of the study are delineated below:

- *Public panic and crisis interactions surged immediately after the shootings, however, the sudden momentum fades away rapidly over time. This was evident from the abrupt reduction in the tweets containing the event-specified words following the Nova Scotia shootings.*
- *While public interactions on recent shootings emerge and disappear quickly, however, discussions on the lessons learned from the historical shootings sustain over a longer period. Discussion on past shooting events dates back to the 1970 Kent State shooting.*
- *The networks of word bigrams revealed highly central keywords such as school and guns. School shootings and their impact on the mental health of the children showed prevalence. While discussing other types of shooting events, people showed more concern about frequently occurring school shootings.*
- *Key interactions include bailing out of shooters, shooting investigation, police response, gun violence, lessons learned from the previous school (Sandy Hook) and mass shootings (El Paso), vehicle ramming (Toronto Van Attack), mobility issues and health concerns during the pandemic, changes in economy and education systems.*
- *Discussions on the challenges due to COVID-19 include topics such as concern about the additional challenge experienced by first responders to accommodate response to active shootings in times of pandemic.*
- *Societal intolerance like racism and political bullying also coappeared with active shooting-related conversations.*

The infographics developed through various data-driven approaches covered both the intermittent and recurrent patterns of risk communication behavior during active shootings. Conversation topics and the most frequent words in tweets derived from the analysis showed the emergence of manifold aspects related to active shooting onsets. Even though the study focused on collecting shooting event-oriented data, it revealed the overall crisis scenario in the Twitter platform. This study would allow first responders and emergency management officials to enhance the capacity of social sharing platforms and facilitate risk communication in no-notice scenarios. Additionally, it can help diverse stakeholders to set up countermeasures for future multi-hazard events.

7. Recommendations for future studies

Due to limited information regarding the origin of tweets, the study could not capture spatial variations of social media interactions on active shootings. A combination of both temporal and spatial differences may capture a more detailed context analysis of crisis events. Future studies could explore how social media can be used to study evacuating or sheltering behavior as well as identifying or avoiding active shooter locations to serve as a decision support tool both for law enforcement and emergency managers. Future research could also try to connect virtual movements of ideas and information (across time and space) and the actual incidents as to behaviors of shooters, affected individuals, and bystanders. Researchers are also recommended to find a better way to establish these connections between actual and virtual environments, to use this information to better inform both real-time operational situational awareness but also operations of responders and traffic managers, to relate this to the sequencing of when and where key events transpired and how it progressed through detection, alert, warning, response, intervention, and recovery.

CRedit authorship contribution statement

M. Ahmed Rusho: Conceptualization, Resources, Methodology, Data curation, Formal analysis, Writing - review & editing, Visualization. **Md Ashraf Ahmed:** Resources, Methodology, Data curation, Formal analysis, Writing - review & editing, Visualization. **Arif**

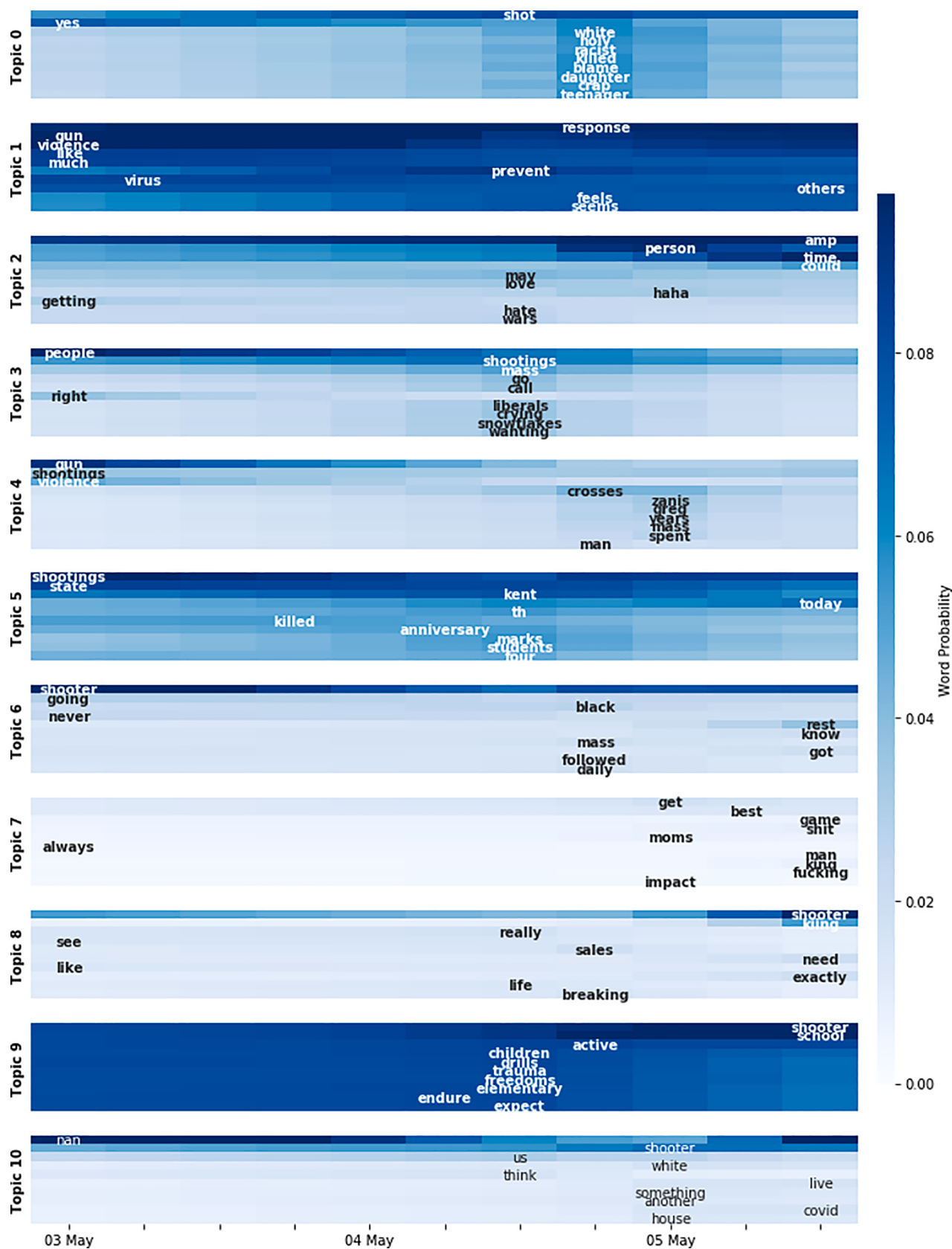


Fig. 5. Word probability in topics over time.

Mohaimin Sadri: Conceptualization, Methodology, Formal analysis, Supervision, Writing - review & editing.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. IIS-2027360. However, the authors are solely responsible for the findings presented in this study.

Author contribution statement

The authors confirm that all authors contributed to the paper as follows: study conception and design; data collection; analysis and interpretation of results; and draft manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

References

- Acar, A., Muraki, Y., 2011. Twitter for crisis communication: Lessons learned from Japan's tsunami disaster. *Int. J. Web Based Commun.* 7, 392–402. <https://doi.org/10.1504/IJWBC.2011.041206>.
- Ahmed, M.A., Sadri, A.M., Pradhananga, P., et al., 2020. Social Media Communication Patterns of Construction Industry in Major Disasters. *Construction Research Congress*.
- Asr, F.T., Taboada, M., 2018. The data challenge in misinformation detection: Source reputation vs. content veracity. In *Proceedings of the First Workshop on Fact Extraction and VERification (FEVER)*.
- Blakinger, K., 2020; Police Killed A Man On His Knees. His Father Wants Answers. | The Marshall Project. <https://www.themarshallproject.org/2020/06/19/it-was-an-execution-nicolas-chavez-was-on-his-knees-when-police-killed-him-his-father-wants-answers>. (accessed 13 Jul 2020d).
- Blei, D., Carin, L., Dunson, D., 2010. Probabilistic topic models. *IEEE Signal Process Mag.* 27, 55–65. <https://doi.org/10.1109/MSP.2010.938079>.
- Bunker, D., Mirbabaie, M., Stieglitz, S., 2017. Convergence Behaviour of Bystanders: An Analysis of 2016 Munich Shooting Twitter Crisis Communication. In: *Proc 28th Australas Conf Inf Syst*, pp. 1–11.
- Caputo P (2005) The Kent State Shootings, 35 Years Later.
- Carrega, C., Zarrell, M., Martinez, L., 2019. 4 dead including suspect after active shooter incident at Naval Air Station Pensacola, police say. *ABC News*.
- Collins, S., 2020. What we know about a mass shooting in Nova Scotia. Canada, Vox. Dictionary. <https://www.dictionary.com/e/slang/snowflake/>.
- Eligon, J., 2019. The El Paso Screed, and the Racist Doctrine Behind It. *N.Y. Times*.
- Engineering, C., 2019. *Comput. Civ. Eng.* 439, 439–446.
- GVA, 2021, Gun violence archive; Mass Shootings in 2020. <https://www.gunviolencearchive.org/reports/mass-shooting?page=7>. (accessed 13 Jul 2020b).
- Herovic, E., Sellnow, T.L., Sellnow, D.D., 2020. Challenges and opportunities for pre-crisis emergency risk communication: lessons learned from the earthquake community. *J. Risk Res.* 23, 349–364. <https://doi.org/10.1080/13669877.2019.1569097>.
- Hu, M., Liu, S., Wei, F., et al., 2012. Breaking news on Twitter. In: *Conf Hum Factors Comput. Syst. – Proc.* <https://doi.org/10.1145/2207676.2208672>.
- Kumar, P., 2017; An Introduction to N-grams: What Are They and Why Do We Need Them? <https://blog.xrds.acm.org/2017/10/introduction-n-grams-need/>, (accessed 13 Jul 2020c).
- Lau, J.H., Collier, N., Baldwin, T., 2012. On-line trend analysis with topic models: Twitter trends detection topic model online. 24th Int Conf Comput Linguist - Proc COLING 2012 Tech Pap 2:1519–1534.
- Levinson, J., 2020; Portland Shootings Increase Despite Stay At Home Order. News | OPB. <https://www.opb.org/news/article/portland-shootings-increase-despite-stay-at-home-order/>. (accessed 13 Jul 2020e).
- Li, L.T., Yang, S., Kavanaugh, A., et al., 2011. Twitter use during an emergency event: The case of the UT Austin shooting. *ACM Int. Conf. Proc. Ser.* 335–336 <https://doi.org/10.1145/2037556.2037613>.
- Liu, B.F., Austin, L., Jin, Y., 2011. How publics respond to crisis communication strategies: The interplay of information form and source. *Public Relat. Rev.* 37, 345–353. <https://doi.org/10.1016/j.pubrev.2011.08.004>.
- Mak, A.K.Y., Song, A.O., 2019. Revisiting social-mediated crisis communication model: The Lancôme regenerative crisis after the Hong Kong Umbrella Movement. *Public Relat. Rev.* 45 <https://doi.org/10.1016/j.pubrev.2019.101812>.
- Martin, S., Liermann, J., Ney, H., 1998. Algorithms for bigram and trigram word clustering. *Speech Commun.* 24, 19–37. [https://doi.org/10.1016/S0167-6393\(97\)00062-9](https://doi.org/10.1016/S0167-6393(97)00062-9).
- Mazer, J.P., Thompson, B., Cherry, J., et al., 2015. Communication in the face of a school crisis: Examining the volume and content of social media mentions during active shooter incidents. *Comput. Human Behav.* 53, 238–248. <https://doi.org/10.1016/j.chb.2015.06.040>.
- Roy, K.C., Hasan, S., Sadri, A.M., Cebrian, M., 2020. Understanding the efficiency of social media based crisis communication during hurricane Sandy. *Int. J. Inf. Manage.* 52, 102060 <https://doi.org/10.1016/j.ijinfomgt.2019.102060>.
- Roy, K.C., Ahmed, M.A., Hasan, S., Sadri, A.M., 2020a. Dynamics of Crisis Communications in Social Media: Spatio-temporal and Text-based Comparative Analyses of Twitter Data from Hurricanes Irma and Michael. *Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM)*.
- Sadri, A.M., Ukkusuri, S.V., Gladwin, H., 2017. The Role of Social Networks and Information Sources on Hurricane Evacuation Decision Making. *Nat. Hazards Rev.* 18 [https://doi.org/10.1061/\(ASCE\)NH.1527-6996.0000244](https://doi.org/10.1061/(ASCE)NH.1527-6996.0000244).
- Sadri, A.M., Hasan, S., Ukkusuri, S.V., Lopez, J.E.S., 2017a. In: *Analyzing Social Interaction Networks from Twitter for Planned Special Events*, pp. 1–20.
- Sadri, A.M., Hasan, S., Ukkusuri, S.V., Cebrian, M., 2018. Crisis communication patterns in social media during hurricane sandy. *Transp. Res. Rec.* 2672, 125–137. <https://doi.org/10.1177/0361198118773896>.
- Scinto, R., 2012. *Sandy Hook Elementary: Newtown*. Oakl Press, Connecticut shooting timeline.
- Silverstein, J., 2020. There were more mass shootings than days in 2019, according to Gun Violence Archive - CBS News. <https://www.cbsnews.com/news/mass-shootings-2019-more-than-days-365/>. (accessed 13 Mar 2020a).
- Tait, M., 2018. Toronto van attack: What we know about the lives lost, the suspect and more. *Globe Mail*.
- Ukkusuri, S.V., Zhan, X., Sadri, A.M., Ye, Q., 2014. Use of social media data to explore crisis informatics study of 2013 Oklahoma Tornado. *Transp. Res. Rec.* 2459, 110–118. <https://doi.org/10.3141/2459-13>.