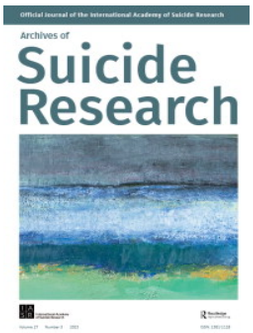

Archives of Suicide Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/usui20

Unlocking the Narrative: Using Text Mining to Reveal the Hidden Factors behind Suicide Related Traffic Crashes

Subasish Das, Boniphace Kutela & Nikhil Menon

To cite this article: Subasish Das, Boniphace Kutela & Nikhil Menon (2023): Unlocking the Narrative: Using Text Mining to Reveal the Hidden Factors behind Suicide Related Traffic Crashes, Archives of Suicide Research, DOI: 10.1080/13811118.2023.2247026

To link to this article: https://doi.org/10.1080/13811118.2023.2247026

Published online: 14 Aug 2023.

ARCHIVES OF SUICIDE RESEARCH
https://doi.org/10.1080/13811118.2023.2247026

Routledge
Taylor & Francis Group

Unlocking the Narrative: Using Text Mining to Reveal the Hidden Factors behind Suicide Related Traffic Crashes

Subasish Das, Boniphace Kutela, and Nikhil Menon

ABSTRACT

Suicide is the deliberate act of ending a person's own life due to multifarious reasons. In the U.S., suicide is the 10th major cause of death. Nearly 45,000 people died by suicide in 2016 across the nation. It is anticipated that not all traffic crashes can be considered as accidents. Traffic crash related injuries are occasionally considered a means of suicide, and some crashes occur due to the suicidal attempts. These attempts can be made by operators of motor vehicles, jumpers into the pathway of trains, and pedestrians deliberately jumping into the vehicle trajectory. There are a handful of studies that have focused on traffic crashes (both railroad and roadway) related to suicidal incidents. This study aimed to explore the insights associated with suicide related traffic crashes (SRTCs) by collecting traffic data for seven years (2010–2016) from Louisiana. At first, exploratory data analysis was performed to examine the five Ws (who, what, why, when, and where) associated with SRTCs. Later, this study applied text network analysis, which was not performed in any of the previous studies, to provide additional contexts of these crashes. The findings of this study can shed lights on an unexplored arena of transportation safety research.

KEYWORDS

Crash narrative analysis; suicide; text mining; text network analysis; traffic crashes

INTRODUCTION

Suicide is considered a social occurrence, and it has progressively increased with the degree of civilization. In traffic crashes, an essential role, apart from external conditions, may be associated with psychological factors. Imprudence is accused as the main cause of crashes, and it implies nervous, emotive, or impulsive subjects that are deprived of the capacity of an efficient and prompt decision.

For over a century, suicidology has explored the relationship between motor vehicle crashes and suicidal behavior from various perspectives, including psychiatry, psychology, and sociology. Some of the pioneers in this field include Menninger (1938), Tabachnik (1973) and Shneidman (1985). They recognized the potential for motor vehicle crashes to be used as a method for suicide and highlighted the need for further research in this area. Menninger (1938) built on the hypotheses of Sigmund Freud regarding a "death instinct" and posited that suicidal behavior results from three major wishes: wish to kill, wish to be killed, and wish to die. Henry and Short (1954) discussed different determinants of suicide. They uncovered a positive relationship between suicide and social status, a negative relationship between suicide and relational system strength, and a negative relationship between suicide and strength of external restraint

of behavior. The authors also discussed how psychological legitimization relates back to how much an outward expression of aggression would affect the flow of love and acceptance for one's experiences. Unnithan (1994) put forth the idea that suicide and homicide are both extreme forms of aggression, just expressed via alternate forms. In a slight departure from the studies focusing on suicide's relation to aggression, Shneidman (1985) posited that depression has an unspoken subliminal suicidal aspect to it due to it having both a self-abnegating and a wish-for-escape component. Farberow (1980) also contributed significantly to the field by introducing the concept of "suicidal intent" in relation to motor vehicle crashes. He argued that certain types of motor vehicle accidents, such as single-car crashes, were more likely to be related to suicide than others. By examining the circumstances surrounding these crashes, Farberow (1980) was able to identify patterns and risk factors associated with suicide-related motor vehicle crashes.

Despite the growing body of literature on this topic, suicide-related traffic crashes remain a complex and under-studied phenomenon. In recent years, advances in data collection and analysis have opened up new avenues for research in this area. By building on the work of earlier researchers and incorporating new techniques, we hope to gain a deeper understanding of the factors that contribute to suicide-related traffic crashes and develop effective prevention strategies.

Transportation injuries have been widely reported as a means of suicide. This includes jumping into the pathway of trains and operators of motor vehicles, but few reports exist on pedestrians intentionally jumping into the pathway of a motor vehicle. Suicide attempts and suicides can have significant effects on other highway authorities, emergency services, and road users. With the lack of research on this issue, data and awareness are generally deficient. Studies need to examine whether suicidal attempts significantly affect overall roadway safety. Driver suicides possess the nomenclature to be disguised as traffic crashes, raising the possibility that suicidal actions are far more widespread than previously thought, with the potential to impact health, insurance companies, and road safety.

The aim of the analysis is to uncover the latent trends in suicide-related traffic crashes. To mitigate the current research gap in understanding suicidal attempts in traffic crashes, this study aimed to design a study by collecting suicide-related traffic crashes from Louisiana. Seven years (2010–2016) of crash data have been collected to identify the required dataset. This study conducted exploratory data analysis and text mining to provide latent insights on the five Ws (who, what, why, when, and where) of these crashes. As the current data analysis procedure acquired police narrated crash reports, the most appropriate data-driven analysis is to perform text mining for extracting insights from unstructured crash reports. The findings from the text mining provide additional contexts to suicide-related traffic crashes.

LITERATURE REVIEW

Studies such as Mann et al. (2005) have demonstrated the value of using rigorous methods to gather relevant publications, extract data based on primary outcomes of interest, and conduct a thorough review of all extracted papers to determine their relevance for

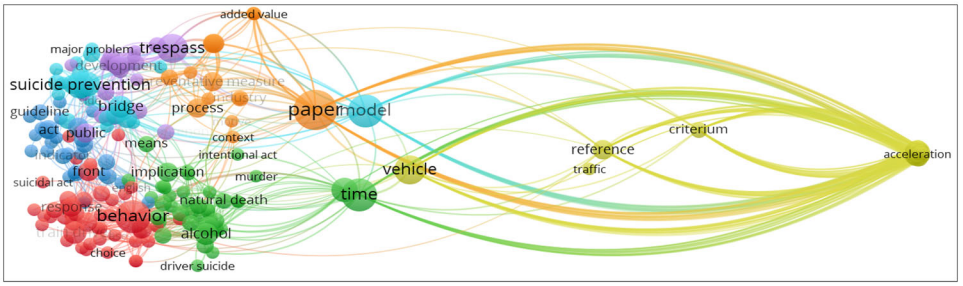

FIGURE 1. Network plot showing the co-occurrence of keywords.

inclusion. For this study, relevant studies associated with suicide-related roadway traffic crashes were collected from the Transportation Research Board's (TRB) Transportation Research Information Services (TRIS) Database, known as TRID (https://trid.trb.org/). This comprehensive search enabled us to identify and gather pertinent literature that sheds light on the complex phenomenon of suicide-related traffic crashes. After removing all irrelevant studies, the current study selected 17 final studies for analysis. To gain a better understanding of the research trends in these studies, a co-occurrence network plot was developed (see Figure 1). This plot helps to visualize the connections between key themes and concepts in the literature and can provide valuable insights into the complex factors that contribute to suicide-related traffic crashes. Some of the words with larger nodes are vehicle, time, acceleration, trespass, bridge, behavior, suicide prevention, and alcohol. At least seven major trends can be visible from the network plot by observing the colors of the lines.

Thoresen and Mehlum (2004) conducted a study on alcohol-related fatal crashes among veterans of military operations, which revealed an increased risk of fatal accidents. Notably, these accidents shared common features with suicide cases, including depression, substance abuse, and social problems. In a separate study, Thoresen et al. (2006) investigated risk factors for suicide in peacekeeping veterans, revealing that involuntary repatriation from peacekeeping service, negative life events prior to deployment, and marital status were all factors linked to an increased risk of suicide. These findings highlight the importance of addressing the mental health needs of veterans, particularly those who have experienced traumatic events in the line of duty, to prevent alcohol-related fatal crashes and suicide.

In a 1964 study, MacDonald (1964) revealed that official statistics are insufficient to determine the extent of deliberate death on highways. Similar findings were also reported in another study (Gussen, 1971). Tabachnik (1973) examined the presence of self-destructive psychological features, suicide-like factors, and suicide to an increased degree in drivers involved in traffic crashes. Damian and Rascanu (1973) psychologically tested 12 neurotic and psychopathic subjects with a questionnaire consisting of 20 questions concerning the "the need of evasion," "tendency to risk," and "non-conformism." Pandurovic, Uzelac, and Jojanovic (1980) identified the motifs for suicide and the injury effects caused by traffic crashes through the presentation of four suicide cases in Belgrade.

Murray and De Leo (2007) used a large-scale community survey to examine suicidal plans and attempts by the drivers. A prospective cohort study conducted by Martiniuk et al. (2010) assessed the risk of deliberate self-harm on traffic crashes among young

drivers. During the follow-up period of 2 years, survey data from recently licensed drivers aged 17–24 years was linked to licensing attempts and police-reported motor vehicle crashes data. Wyatt, Squires, Collis, and Broadley (2009) conducted a prospective observational study of fatal road traffic collisions in south-east Scotland. This study showed that 2.8% of all road fatalities occurred due to suicidal attempts. Sansone, Lam, and Wiederman (2010) examined the relationship of prior suicidal attempts with a history of reckless driving.

Some studies explored systemic analysis to determine the association between the intent of suicide and single-car crash drivers. Pompili et al. (2012) shed light on overlooked aspects that may play a crucial role in reducing the number of traffic crash fatalities by reviewing relevant literature. About 1% of all suicidal attempts used in Switzerland are the result of road traffic crashes, although a higher rate may be indicated by unclassifiable cases. Using the suicide database investigated by Swiss institutes of forensic medicine from 2000–2010, Gauthier, Reisch, Ajdacic-Gross, and Bartsch (2015) conducted a retrospective data analysis of road traffic suicides with a standardized assessment sheet. Of the 4,885 suicides in Switzerland, 76 cases (1.56%) were identified as road traffic suicide. Harrison (2017) revealed how vehicles, roads, and road infrastructure are being used by individuals in the UK with the intent of ending their lives. This study estimated that there are likely to be approximately 50 suicides each year on UK roads. Several new studies focus on railway-related suicidal attempts. Reynders et al. (2011) aimed to compare railway suicides from two data sources to assess the reliability of national suicide data. Results showed that the European Railway Agency (ERA) registered significantly higher railway suicides compared to national data in six out of fifteen countries, with Italy and France recording more railway-related fatalities in ERA statistics than national statistical offices. The study highlights the need for caution when comparing national suicide rates and calls for standardization of death registration procedures at the European level.

As seen from the earlier efforts described here, while the share of crashes that are directly attributable to suicides or self-harm may be less in comparison to other causes, there is possible under-representation in obtaining the right amount of these crashes due to the challenges associated with data collection and categorization of crashes that are recorded into the various databases. To achieve better results in categorizing crashes as well as overcoming the challenges with data collection, there has been a push for using methods that provide additional latent information that provides more insight into driver behavior, their state of mind, and potential triggers that may have contributed to the crash. In this regard, crash narrative analysis has been becoming popular among the researchers due to the power of revealing the latent information from unstructured textual contents. The literature review shows that no studies explored suicide-related crashes by exploring crash narrative analysis. This study aims to mitigate this research gap by exploring the unstructured texts and structured crash data associated with suicidal attempts.

METHODOLOGY

Data Collection

The data collection process for this study involved collecting redacted crash narrative reports for a period of seven years, from 2010 to 2016. The research team used two

TABLE 1. Examples of crash narrative excerpts in the redacted crash narrative reports.

No.	Year	Suicide related texts in Crash Narrative Report
1	2010	driver one advised he was traveling. driver one stated he purposely drove vehicle one into a tree because he attempted to commit suicide.
2	2010	driver1, who was identified as xxx, restrained due to him being combative and threatening to commit suicide. he was a friend of driver1 and advised that driver1 had mental/psychological disorder(s).
3	2010	was unable to determine if driver of vehicle one suffer any injuries related to the crash, due to the apparent suicide just after the crash occurred. i did detect the distinct odor of an alcoholic beverage inside the passenger compartment of vehicle one.
4	2010	she and her boyfriend had an argument, so she tried to hit a tree to commit suicide.
5	2010	she had frequent black outs due to her condition and should not be driving. she stated she has attempted suicide several times.

Note. The excerpts are generated from the redacted crash narrative reports using text mining; spelling, grammatical errors, and punctuations are not corrected in these excerpts.

keywords, "suicide" and "suicidal," to identify relevant reports, resulting in an initial dataset of 80 crash reports. To refine the dataset, a manual effort was conducted to ensure that only reports related to suicide-related traffic crashes were included. This effort revealed that some reports contained references to "suicide lane" and "suicide door," which were not related to suicide-related traffic crashes. These reports were removed from the dataset. Ultimately, the final dataset consisted of 71 crash reports that were related to suicide-related traffic crashes. While the data collection process may have had limitations, including the possibility of missing relevant reports or inadvertently excluding relevant information, every effort was made to ensure the accuracy and validity of the final dataset.

Crash Narrative Excerpts

Table 1 provides some examples of the reporting patterns of suicide-related crashes. Using text processing tools, excerpts with "suicide" or "suicidal" (report location where any of these two words were present) and the neighboring sentences were produced. The excerpts are reproduced here to give the interested readers a broad overview of the contexts associated with police reported crash narratives. It is found that majority of the texts in the crash narrative report discuss engaging in self-destructive behavior immediately prior to the occurrence of the crash. Many of the drivers who were exhibiting suicidal behavior, or were under advisement to not be driving due to their existing conditions, were behind the wheel at the time of the crash. It is also telling how there were instances where drivers, due to their fragile state of mind, engaged in not only self-harming themselves but also put their passengers at risk due to arguments or disagreements between the two parties. In a minority of cases, an external agent such as alcohol or other intoxicants were found in blood samples of the drivers in addition to their seemingly fragile state of mind at the time of the crash. These external factors have positive effects and could have contributed to the potential lapse in judgment leading to these crashes.

Exploratory Data Analysis

Table 2 lists the distribution of the variable categories of the crashes associated with suicidal attempts. Most of the crashes do not involve alcohol impairments (a little over

TABLE 2. Percentage distribution of the key attributes.

Variable	Attribute	Percent	Variable	Attribute	Percent
Alcohol presence	No Alcohol/Drugs	71.31	Collision type	Single Vehicle	35.25
	Yes (Alcohol)	6.56		Rear End	15.57
	Yes (Alcohol and Drugs)	4.10		Right Angle	12.30
	Not reported	2.46		Sideswipe - Same Di.	11.48
	Yes (Drugs Present)	0.82		Sideswipe - Opposite Dir.	4.10
	Unknown	14.75		Head-On	3.28
Driver injury	No Injury	54.92		Left Turn - Angle	3.28
	Possible/Complaint	31.15		Left Turn - Opposite Dir.	1.64
	Non-Incapacitating/Moderate	8.20		Right Turn - Same Dir.	1.64
	Incapacitating/Severe	2.46		Other	11.48
	Not reported	3.28	Reason for movement	Normal Movement	32.79
Driver race	White	67.21		Due to Driver Violation	30.33
	African American	27.05		Due to Driver Condition	15.57
	Asian	0.82		Other	9.02
	Not reported	4.92		For Traffic Control	0.82
Driver gender	Male	54.92		To Avoid Pedestrian	0.82
	Female	40.16		Vehicle out of Control, Not Passing	0.82
	Not reported	4.92		Unknown	8.20
Hit and run	No	79.51		Not reported	1.64
	Yes	20.49	Vehicles involved	Two	54.10
Highway type	City Street	32.79		One	40.98
	Parish Road	27.05		Three	4.92
	State Hwy	27.05	Primary contribution factor	Violations	63.93
	U.S. Hwy	7.38		Movement Prior To Crash	14.75
	Interstate	4.92		Condition of Driver	12.30
	Not reported	0.82		Pedestrian Actions	5.74
Intersection	No	65.57		Condition of Pedestrian	3.28
	Yes	34.43	Crash severity	No Injury	40.16
Lighting condition	Daylight	55.74		Complaint	36.07
	Dark - Cont. Street Lt	22.13		Moderate	15.57
	Dark - No Street Lt	11.48		Severe	8.20
	Dark - Street Light at Int.	4.10	Vehicle lighting condition	Headlights On	36.89
	Dawn	2.46		Unknown	28.69
	Not reported	1.64		Headlights Off	23.77
	Dusk	0.82		Daytime Running Lights	10.66
	Other	1.64	Violation type	No Violations	38.52
Locality type	Business Continuous	35.25		Careless Operation	31.97
	Business, Mixed Residential	25.41		Driver Condition	5.74
	Residential District	21.31		Failure to Yield	5.74
	Residential Scattered	8.20		Cutting In, Improper Passing	2.46
	Open Country	6.56		Disregarded Traffic Control	0.82
	Manufacturing or Industrial	0.82		Exceeding Stated Speed Limit	0.82
	Other	2.46		Following too Closely	0.82
				Unknown	4.10
				Other	9.02

10% of the crashes are associated with impairment). Around 11% of the crashes are associated with severe or moderate injuries. White Americans are more involved in these crashes compared to African Americans and other minorities. A larger majority of the drivers involved in these crashes were males (54%). Close to 60% of the crashes

occurred on city streets or parish roads, with a significant majority of crashes occurring at non-intersections (66%). It is interesting to note that a significant majority of crashes (greater than 60%) occurred in conditions with ample lighting (either daylight or dark with continuous street lighting). Business localities (with or without residential mixes) constitute more than 60% of the crash locations. This could be potentially likely due to the higher speeds experienced along these corridors in comparison to residential corridors.

Single vehicle crashes show the highest proportion of collision type (35%), which is perhaps indicative of the self-destructive mindset of the driver leading to the crash. While almost one-third of the crashes (33%) were caused in normal movement circumstances, it is also worthwhile to notice that an additional one-sixth of the crashes (16%) occurred due to the condition of the driver. While there were multiple vehicles involved in most of the crashes, it is important to point out that close to 40% of the crashes only involved a single vehicle. This is perhaps indicative of the fragile state of the driver, potentially due to suicidal behavior and other agents (as discussed previously) leading to the eventual crash. A significant share of crashes occurred due to violations (64%) while about one in every eight crashes was attributed to the condition of the driver (12%). Crash severity percentages are also provided to give an overall picture of injuries associated with both parties if the crash was involved with multiple vehicles. The proportions by injury types are not widely different than driver injury types. Vehicle headlights were turned on in around 35% of the crashes. However, in many cases, this information is unknown (around 19%). Correlating this statistic with the percentage of crashes that occurred during lighting conditions, it is intuitive why the percentages of headlights turned on were just over one-third of all crashes. Lastly, careless operation represents almost one-third of the crashes (32%) of the crashes in addition to driver condition (6%). Interestingly, less than 1% of the crashes occurred due to speed violations, although this may be slightly under-represented owing to the challenges associated with data collection and the associated characterization of crashes after the event.

Text Network Analysis

Crash narrative text network analysis is a method used to analyze the narrative structure of crash reports and identify the key themes and connections between them in order to gain a deeper understanding of the circumstances and factors that contribute to suicide-related traffic crashes. There are several reasons why crash narrative text network analysis is a valuable tool in this context. Firstly, suicide-related traffic crashes are complex events that involve multiple factors and influences. By breaking down the narrative structure of crash reports and identifying key themes and connections, researchers can gain a more detailed understanding of the specific factors that contribute to these events. Secondly, crash narrative text network analysis can help to identify patterns and trends in suicide-related traffic crashes. By examining multiple crash reports and mapping out the connections between them, researchers can identify common themes and factors that contribute to these events. This can inform the development of targeted prevention strategies and interventions. Finally, crash narrative text network analysis can help to identify gaps in current knowledge and highlight areas for further research.

By analyzing the narrative structure of crash reports, researchers can identify themes and connections that may not have been previously recognized, leading to new insights and avenues for future investigation.

Text network analysis uses nodes, which represents keywords and links, which represents co-occurrence of the keywords to develop insights from unstructured textual data. Although this approach is relatively new, it has been applied in several aspects of transportation engineering such as bike-sharing, dynamic message signs operations, and explaining the relationship between transportation research experts (Kutela et al., 2021; Kutela & Teng, 2021). In the context of suicide-related traffic crashes, text network analysis can be a valuable tool for understanding the factors and circumstances that contribute to these events. By analyzing the narrative structure of crash reports and identifying key themes and connections, researchers can gain insights into the specific factors that contribute to suicide-related traffic crashes. For example, they may identify common themes such as mental health issues, substance abuse, or relationship problems, and map out the connections between these factors and the events leading up to the crash.

Basically, the analysis involved two major steps: building the network and interpreting it. To build a network, text normalization (removal of stop words, setting all texts to lower case), transformation from unstructured to structured data (matrix formulation), and generation of nodes and links should be performed. The first two procedures are commonly performed for most text mining approaches, but the network creation procedure is worth explaining. In this step, the matrix of keyword formulated in the transformation step is used. The algorithm searches for a first pair of keywords and plots it to the network. The algorithm then continues to search for other pairs of keywords. If the next pair has a keyword that is already mapped, the algorithm adds the weight to the mapped and new keywords to form the second pair. If the pair is new, the algorithm maps it as a new pair with the associated weights of co-occurrence.

The last step is the extraction of insights used for in-depth interpretation of the network. In this step, several approached can be used, however, in this study the keyword frequency, communities, degree centrality, and co-occurred keywords are used. Communities are formed by keywords with similar themes. On the other hand, the collocation analysis focuses on the keywords that appear next to each other in a sentence. Such keywords provide deeper insights than co-occurring keywords. Lastly, since the analysis normally results in large networks, only the top 40 keywords were used for interpretations in this study.

RESULTS AND DISCUSSIONS

Before analyzing the textual data, further cleaning of the data was necessary to ensure that the text is uniform. The preliminary review of the text showed that the driver of the first vehicle could be written as *driver 1, driver#1, driver number 1, driver number one,* etc. Therefore, all these nomenclatures were converted to *driver-1.* A similar process was done for the driver of the second vehicle.

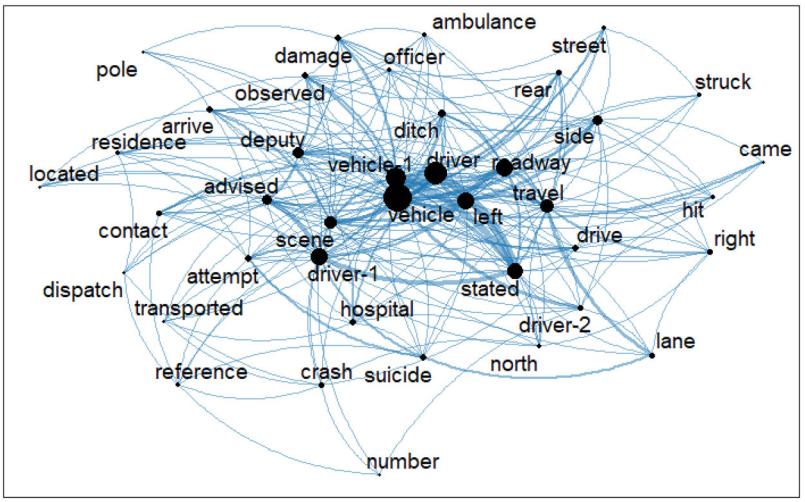

FIGURE 2. Overall text network for victims (25 years old or younger).

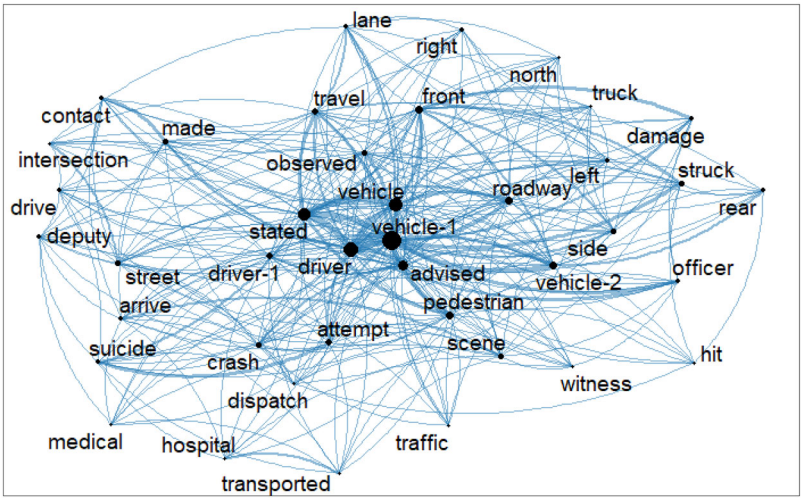

FIGURE 3. Overall text network for victims (over 25 years old).

Figures 2 and 3 presents the overall text networks and driver/officer excerpts of the narrative for victims with 25 years or less and over 25 years, respectively. Several similarities and differences are portrayed in the Figures.

Figure 2 presents the text network for victims who are 25 years old or younger. The text network is heavily centered on five large nodes with the keywords: *vehicle-1*, *driver*, *roadway*, *driver-1*, and *vehicle*. These keywords show the main participants and the locations related to crashes, with *driver* and *vehicle* showing the participants and *roadway* showing the locations of crashes. There are also some collocations by linking keywords in the text network, like *driver* and *stated*, *vehicle* and *driver*, *suicide* and *attempt*, *travel* and *lane*, and *vehicle* and *damage*, which also illustrates the relationship between the participants aged 25 or younger and the locations related to crashes.

Figure 3 presents a similar text network for victims aged over 25 years old. The text network is heavily centered on five large nodes with the keywords: *vehicle-1*, *driver*, *roadway*, *stated*, and *pedestrian*. These keywords show the main participants and the information related to crashes like *driver*, *vehicle,* and *pedestrian* showing the participants and *stated* showing the delivery method of the information. Some collocations by linking keywords are shown in the text network, like *driver* and *stated*, *vehicle* and *driver*, *suicide* and *attempt*, *vehicle* and *stated*, and *vehicle* and *pedestrian*, which also illustrates the relationship between the participants aged over 25 and the information delivery method related to crashes.

Both Figures 2 and 3 show the key participants in the crash, the source of the information, police officer involvements, medical care provision, the location of the crash, and others. The text networks are centered on five major keywords, which are *vehicle-1, vehicle, driver, and stated.* These keywords and the connections between them show the key participants (*driver* and *driver of vehicle 1*) and how the information was delivered (*stated*). Obviously, in the suicide attempts, the sources of crash narratives are normally the drivers involved in the crashes unless said driver dies in a crash, in which case witnesses, if available, provide the crash information.

Additionally, more key participants, *pedestrian* and *driver-1,* are observed in the networks. However, *driver-1* is more dominant in the text network for 25 years or less, while pedestrians are more vivid for the network for victims with more than 25 years. The implication is that the proportion of crashes that involved pedestrians committing suicide is higher for victims with more than 25 years than for the counterpart. On the other hand, the keyword *driver-1* is typically utilized to indicate the first driver, who most of the time is at fault. Furthermore, both networks have keywords such as *struck, rear, front, right,* and *truck* that show the details of the impact type and vehicle used. Also, the networks show the victim's intention (suicide attempt), the medical care provided to the victim, and the involvement of the witnesses and police officers (deputy/officer) at the crash scene.

Although Figures 2 and 3 provided great details of the crash, more useful details are presented in Figures 4 and 5, and the top associated metrics are presented in Table 3. These figures were produced by zeroing on the narratives that explain what happened just before the crashes. Such narratives were stated by either the witnesses, survived drivers, or the police officer.

Figure 4 presents the text network for the causes of the incidents for victims who aged 25 years old or younger. There are some large nodes which are more heavily used, like *travel*, *lane*, *roadway*, *side*, and *traffic*. These keywords are stated by who were related to the crashes, including the witnesses, drivers, and police officers. Some collocations were formed by linking keywords are showing in the text network, like *travel* and *lane*, *left* and *right*, *center* and *lane*, *beverage* and *alcoholic*, *driver* and *U-turn*, which also illustrate the reasons and the specific locations of the crashes.

Figure 5 presents the text network for the causes of the incidents for victims who were aged over 25 years old. Figure 5 looks more centered than Figure 4, showing that the causes of incidents for victims are more related. There are some large nodes centered which are more heavily used, like *pedestrian*, *street*, *roadway*, *side*, and *front*. These keywords are also stated by those who were related to the crashes. Some

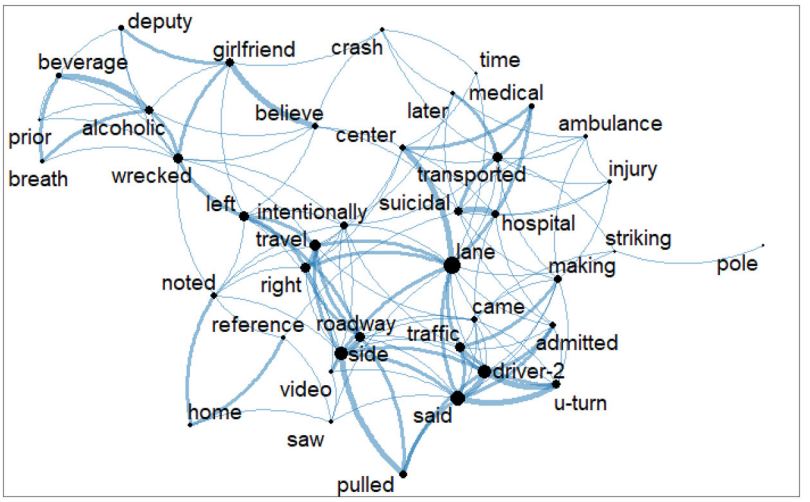

FIGURE 4. Text network for the causes of the incidents for victims (25 years old or younger).

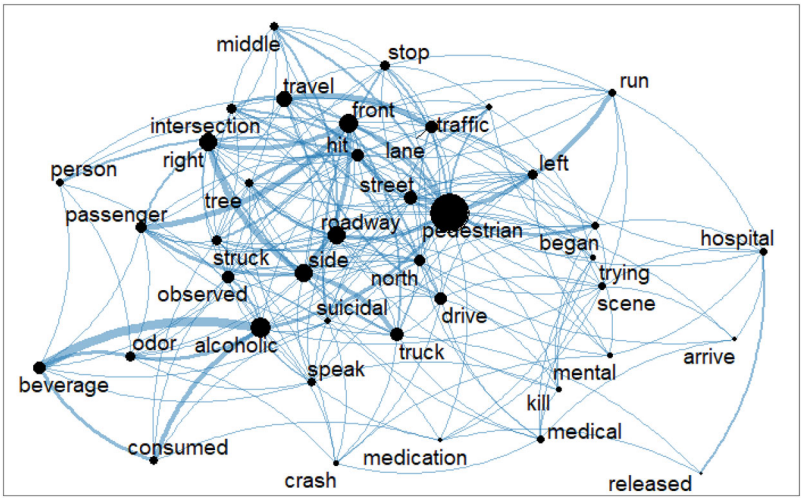

FIGURE 5. Text network for the causes of the incidents for victims (over 25 years old).

collocations formed by linking keywords are shown in the text network, like *front* and *side*, *intersection* and *side*, *street* and *pedestrian*, *beverage* and *alcoholic*, and *front* and *pedestrian*, which also illustrate the reasons and the specific locations of the crashes.

Overall, the networks show that there were two types of suicide attempts: by pedestrians and by drivers. The network shows that suicide attempts involving pedestrians were common for drivers aged over 25 years. This observation can be deduced from the text network for drivers aged over 25 years (Figure 5) and the associated metrics (Table 3).

The network shows that the keyword *pedestrian* has the largest node size, whereby the corresponding network metrics show that the same keyword has the highest frequency and degree centrality score. The implication is that apart from being the most

TABLE 3. Text networks metrics for young and older victims in suicide related traffic crashes.

	Text Network Metrics for Victims (25 years old or younger)						
	Frequency			Co-occurrence		Degree centrality	
Rank	Keyword	Score	Observation	Keyword	Count	Keyword	Score
1	lane	6	3	center lane	3	lane	16
2	hospital	6	5	driver-2 said	3	said	12
3	deputy	5	3	right travel	3	transported	12
4	suicidal	5	5	suicidal hospital	3	intentionally	11
5	said	4	2	believe girlfriend	3	wrecked	11
6	intentionally	4	4	alcoholic beverage	3	driver 2	11
7	girlfriend	4	3	pulled side	3	side	11
8	wrecked	4	4	driver-2 U-turn	3	travel	10
9	driver 2	4	2	roadway side	3	right	9
10	transported	4	3	alcoholic breath	2	traffic	9
	Text Network Metrics for Victims (over 25 years old)						
	Frequency			Co-occurrence		Degree centrality	
Rank	Keyword	Score	Observation	Keyword	Count	Keyword	Score
1	pedestrian	31	7	alcohol beverage	9	pedestrian	27
2	roadway	15	11	passenger front	6	roadway	20
3	truck	12	8	right side	6	street	19
4	travel	12	7	travel lane	6	observed	18
5	crash	11	10	pedestrian run	5	right	17
6	side	11	8	truck side	5	hit	17
7	front	11	8	right front	5	north	17
8	drive	11	6	pedestrian street	5	side	17
9	right	9	6	roadway pedestrian	4	left	16
10	tree	9	7	hit tree	4	front	16

frequent keyword, it is also linked to many other keywords as it is the center of the narratives.

Further, the co-occurring keywords displayed in the networks and in the table for the associated metrics (see Table 3) differ significantly for the two networks. For instance, the co-occurring keywords *travel lane* in Figures 4 and 5 shows that most of the incidents occurred on the travel lanes. However, the text network for victims aged 25 years or less specified the lane that these incidents occurred by using the co-occurred keywords *center lane*. On the other hand, Figure 5 did not specify the lane that most of the incidents occurred. It can further be observed that the police officer used alcohol odor as the first indication of the further search for driving under the influence (DUI), as indicated by the presence of the keyword *alcoholic, beverage, and odor*. This observation was common for both networks; however, for drivers older than 25 years, it was indicated that a significant number of them consumed alcoholic beverages as indicated by the co-occurred keywords *consumed alcoholic beverage*. Additionally, several victims older than 25 years of age had mental conditions associated with their suicidal attempts. This observation is supported by the presence of the keywords *mental* and *medications* in the text network for victims older than 25 years compared to young ones. Furthermore, there were involvements of the girlfriends as either the source of the suicide attempts or the source of information for victims aged 25 years or less. Lastly, the driver-involved suicide attempts involved trucks hitting a tree, especially on the passenger side for older drivers (over 25 years old), while young drivers (less than 25 years old) were more likely to strike a pole.

CONCLUSIONS

This study conducted a unique procedure by collecting suicide related crashes by using a text searching algorithm in crash narrative data reports. For this task, this study used seven years of traffic crash narrative reports (personal information redacted version) from Louisiana. After preparing the data, this study conducted two major tasks: 1) performed exploratory data analysis to identify the five Ws, and 2) applied text network analysis to identify the key trends. This study also provides excerpts of the identified crashes by using a text summarizing approach, which produces sentence excerpts surrounding the keyword by removing all personal information. These excerpts are self-explanatory and can produce additional contexts of this analysis. Overall mental heath improvement can be considered as a countermeasure. Specific intervention design to reduce suicidal traffic crash on roadways is difficult. However, suicidal deaths on railway tracks can be preventable by using countermeasures such as physical barriers. Overall, crash narrative text network analysis is a powerful tool for understanding the complex factors that contribute to suicide-related traffic crashes. By providing a detailed, nuanced analysis of crash reports, this method can help to inform the development of effective prevention strategies and interventions and pave the way for further research in this area.

The current study shows the value of text mining to uncover latent trends in suicide-related traffic crashes. However, this study has certain limitations. First, the sample size is small. It may be argued that this issue is not one of critical concern for transportation safety engineers due to the low number of crashes associated with suicide. However, it should be noted that the real crash causation factors are unknown in many cases. A more in-depth investigation is needed to tie a crash with suicide, which is not performed in many cases. Second, the study is based on the police reported narration. It is anticipated that the numbers of suicide-related crashes are underreported. Crash reports will not be labeled as "suicide-related" if the keywords such as suicide or suicidal are missing in the narrative reports. This study can be considered a starting point to explore this important issue. Future studies can extend the current research by using a larger set of data and applying other suitable analytical tools.

ACKNOWLEDGMENTS

The authors express our sincere gratitude to the anonymous reviewers for their invaluable feedback and constructive suggestions that greatly enhanced the quality and clarity of our manuscript. Their thorough review and insightful comments have undoubtedly strengthened the rigor and impact of this research. The authors sincerely appreciate their time, effort, and expertise in providing valuable guidance throughout the review process.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author(s).

AUTHOR NOTES

Subasish Das, Ph.D. Civil Engineering Program, Texas State University, San Marcos, TX, USA. Boniphace Kutela, Ph.D. Texas A&M Transportation Institute, Texas A&M University System,

College Station, TX, USA. Nikhil Menon, Ph.D. Civil Engineering Department, Pennsylvania State University Harrisburg, Middletown, PA, USA.

Correspondence concerning this article should be addressed to Civil Engineering Program, Texas State University, 601 University Drive, San Marcos, TX 78666, USA. Email: subasish@txstate.edu; s-das@tti.tamu.edu

REFERENCES

Damian, N., & Rascanu, R. (1973). Some Aspects of Suicide Among Motor Car Drivers. *Neurologia, Psihiatria, Neurochirurgua*, *18*(3), 263–268.

Farberow, N. L. (1980). The many faces of suicide. *Annals of the New York Academy of Sciences*, *347*(1), 26–36.

Gauthier, S., Reisch, T., Ajdacic-Gross, V., & Bartsch, C. (2015). Road Traffic Suicide in Switzerland. *Traffic Injury Prevention*, *16*(8), 768–772. doi:10.1080/15389588.2015.1021419

Gussen, J. (1971). Drivers Injured in One-Car Accidents and Survivors of Serious Suicide Attempts. *Accident Analysis & Prevention*, *3*(4), 269–276. doi:10.1016/0001-4575(71)90005-4

Harrison, K. (2017). *Suicide on UK Roads: Lifting the Lid*. London: Parliamentary Advisory Council for Transport Safety

Henry, R., & Short, J. (1954). Suicidal intent in automobile accidents. *American Journal of Psychiatry*, *110*(11), 825–827.

Kutela, B., Langa, N., Mwende, S., Kidando, E., Kitali, A. E., & Bansal, P. (2021). A text mining approach to elicit public perception of bike-sharing systems. *Travel Behaviour and Society*, *24*, 113–123. doi:10.1016/j.tbs.2021.03.002

Kutela, B., & Teng, H. (2021). The Use of Dynamic Message Signs (DMSs) on the Freeways: An Empirical Analysis of DMSs Logs and Survey Data. *Journal of Transportation Technologies*, *11*(01), 90–107. doi:10.4236/jtts.2021.111006

MacDonald, J. M. (1964). Suicide and Homicide by Automobile. *The American Journal of Psychiatry*, *121*(4), 366–370. doi:10.1176/ajp.121.4.366

Mann, J. J., Apter, A., Bertolote, J., Beautrais, A., Currier, D., Haas, A., … Hendin, H. (2005). Suicide Prevention Strategies: A Systematic Review. *JAMA*, *294*(16), 2064–2074. doi:10.1001/jama.294.16.2064

Martiniuk, A. L., Ivers, R. Q., Glozier, N., Patton, G. C., Lam, L. T., Boufous, S., … Norton, R. (2010). Self-Harm and Risk of Motor Vehicle Crashes in Youth: The Drive Prospective Cohort Study. *Injury Prevention*, *16*(Supplement 1), A100–A100. doi:10.1136/ip.2010.029215.360

Menninger, K. (1938). *Man against himself*. San Diego, CA: Harcourt, Brace and Company.

Murray, D., & De Leo, D. (2007). Suicidal Behavior by Motor Vehicle Collision. *Traffic Injury Prevention*, *8*(3), 244–247. doi:10.1080/15389580701329351

Pandurovic, S., Uzelac, Z., & Jojanovic, D. (1980). The Effects of Traffic Accident Injuries as a Motive for Suicide. *Psihomedicina Saobracaja*, *10*(1/2), 7–9.

Pompili, M., Serafini, G., Innamorati, M., Montebovi, F., Palermo, M., Campi, S., … Girardi, P. (2012). Car Accidents as a Method of Suicide: A Comprehensive Overview. *Forensic Science International*, *223*(1-3), 1–9. doi:10.1016/j.forsciint.2012.04.012

Reynders, A., Scheerder, G., & Van Audenhove, C. (2011). The reliability of suicide rates: An analysis of railway suicides from two sources in fifteen European countries. *Journal of Affective Disorders*, *131*(1-3), 120–127. doi:10.1016/j.jad.2010.11.003

Sansone, R. A., Lam, C., & Wiederman, M. W. (2010). History of attempted suicide and reckless driving: A cross-sectional study in primary care. *Primary Care Companion Journal of Clinical Psychiatry*, *12*(4), PCC.09m00888.

Shneidman, E. S. (1985). Some Thoughts on Grief and Mourning. *Suicide & Life-Threatening Behavior*, *15*(1), 51–55. doi:10.1111/j.1943-278x.1985.tb00788.x

Tabachnik, N. (1973). *Accident or suicide?* Destruction by Automobile. Springfield, IL: Charles C Thomas Pub Ltd.

Thoresen, S., & Mehlum, L. (2004). Risk Factors for Fatal Accidents and Suicides in Peacekeepers: Is There an Overlap? *Military Medicine*, *169*(12), 988–993. doi:10.7205/milmed.169.12.988

Thoresen, S., Mehlum, L., Røysamb, E., & Tønnessen, A. (2006). Risk Factors for Completed Suicide in Veterans of Peacekeeping: Repatriation, Negative Life Events, and Marital Status. *Archives of Suicide Research : official Journal of the International Academy for Suicide Research*, *10*(4), 353–363. doi:10.1080/13811110600791106

Unnithan, N. P. (1994). *The currents of lethal violence: An integrated model of suicide and homicide*. New York: State Univ. of New York Press.

Wyatt, J. P., Squires, T., Collis, S., & Broadley, R. (2009). Road Traffic Suicides. *Journal of Forensic and Legal Medicine*, *16*(4), 212–214. doi:10.1016/j.jflm.2008.12.003