Towards a Framework for the Automatic Detection of Crisis Emotions on Social Media: a Corpus Analysis of the Tweets Posted after the Crash of Germanwings Flight 9525

Veronique Hoste
and Cynthia Van Hee

LT3 Language and Translation Technology Team
Ghent University
Ghent, Belgium
Email: firstname.lastname@ugent.be

Karolien Poels

Department of Communication Studies
University of Antwerp
Antwerp, Belgium
Email: karolien.poels@uantwerpen.be

***Abstract*—Social media, and in particular Twitter, are increasingly being utilized during crises. It has been shown that tweets offer valuable real-time information for decision-making. Given the vast amount of data available on the Web, there is a need for intelligent ways to select and retrieve the desired information. Analyzing sentiment and emotions in online text is one option for distinguishing relevant from irrelevant information. In this study, we investigate to what extent automatic sentiment analysis techniques can be used for detecting crisis emotions on Twitter. Therefore, a corpus of tweets posted after the crash of Germanwings Flight 9525 was built and labeled with polarity and emotion information. Preliminary results show better classification results for the negative sentiment class compared to the positive class. An analysis of the more fine-grained emotion classification reveals that sympathy and anger are the most frequently expressed emotions in our corpus. To further enhance the performance of emotion classification in online crisis communication, it is crucial to accurately detect i) the object of the crisis emotion and ii) the characteristics of the sender.**

Keywords–emotion detection; social media; natural language processing; organizational crisis; crisis communication.

I. Introduction

The use of social media has thrived over the past few years. As a consequence, the ways in which people communicate during crisis situations have changed. Especially the microblogging service Twitter has become a very popular web application for seeking and defusing crisis-related information [1], [2], [3]. Furthermore, it is an ideal way for crisis managers to demonstrate their compassion, concern, and empathy to stakeholders in case of an organizational crisis. An organizational crisis can be described as “the perception of an unpredictable event that threatens important expectancies of stakeholders and can seriously impact an organization’s performance and generate negative outcomes” [4]. An event is partially defined as a crisis by the perceptions of stakeholders [4]. Bryson [5] defines a stakeholder as “a person or a group that is influenced by or has an influence on an organization”. Crises interfere with some stakeholder expectancies, which results in people becoming angry and upset. As a consequence, the organization is perceived less positively and its reputation is damaged. It is critical for organizations and public relations practitioners working in the field of crisis communication to have knowledge about how to shape the appropriate strategies in response to crises. Coombs’ [6], [7] Situational Crisis Communication Theory (SCCT) is a dominant theory on crisis response strategies. It takes an audience-centred approach in order to understand stakeholders’ reactions in crisis situations by examining their attribution of crisis responsibility [8]. Attribution theory posits that people will make judgements about the causes of events, especially those that are unexpected and generate negative outcomes [6]. Since crises are (mostly) unforeseen and negative, they are just the type of event that will produce attributions. If stakeholders think an organization should have been able to control a crisis or has made serious mistakes, they will blame the organization for the crisis. Furthermore, greater attributions of responsibility result in stronger feelings of anger and more negative visions on people and organizations [9], something that should be carefully monitored.

All of this shows that understanding people’s reactions and emotions during a crisis is crucial for organizations. In this paper, we explore how sentiment analysis can be used to understand how publics consume crisis information. To this end, a state-of-the-art sentiment analysis system was applied to a Twitter dataset, which we collected after the crash of a Germanwings aircraft in the French Alps in 2015. While sentiment analysis systems classifies the tweets according to their polarity (positive, negative or neutral), they do not give insights into the more fine-grained emotions expressed in texts. In order to better understand the types of emotions expressed in our corpus, we further labeled the data with the crisis-related emotion categories as proposed by Jin et al. [9] and report our findings.

The remainder of this paper is organized as follows: Section II presents a literature overview on the analysis of sentiment and emotions in crisis communication. In Section III, we describe the experiments on sentiment classification and emotion detection whereas Section IV discusses the findings of our analysis. Finally, in Section V we draw some conclusions and present prospects for future work.

II. Sentiment and Emotions in Crisis Communication

In order to handle a crisis effectively, it is crucial for crisis managers to understand how emotions are related to crisis

 ISBN: 978-1-61208-519-7

TABLE I. Occurrence of emotion classes in the gold standard corpus.

Emotion class	# tweets	Example tweet
Anger	25	This documentary about Andreas Lubitz is making my blood boil #GermanWingsCrash
Fear	4	Thanks to the evil #GermanWingsCrash I'm officially scared to fly, they should allow us to talk and meet our pilot incase.
Apprehension	4	If the pilot used an axe on the door, whats to stop a terrorist? What other potential weapons r laying round on flights? #GermanWingsCrash
Confusion	2	Should I be worried or reassured by the #GermanWingsCrash? It is good to know that the doors won't open from the outside...but then again...
Contempt	21	So this guy takes a picture in front of the Golden Gate Bridge..The most used bridge for suicide jumps. Dude why not then? #GermanWingsCrash
Disgust	9	The Daily Mail coverage of the #GermanWingsCrash has been repugnant. Headlines like 'how the nazis led to killer co-pilot' help no one.
Embarrassment	0	-
Guilt	0	-
Sadness	14	I feel really sad for the 150 families who are suffering as a result of the #GermanWingsCrash. Beyond tragic.
Surprise	1	Blown away. Pilot locked out of the #Germanwings cockpit?!? I thought I heard it all. #German-WingsCrash
Sympathy	26	Our thoughts and prayers go out to those who lost loved ones in the #GermanWingsCrash May God be with you in these hard times.
Other	2	I'm thinking this attn on #AndreasLubitz and the #GermanWingsCrash is overdone. It's tragic & I would rather see the focus on the victims.

responsibility and crisis communication strategies. Therefore, crisis managers should understand how crisis situations are appraised and evaluated by stakeholders [8]. It was found that stronger attributions of crisis responsibility result in feelings of anger and in some extreme cases in *schadenfreude* (i.e., getting pleasure from the pain of others) toward the organization [10]. Moreover, feelings of sympathy for the organization reduce if a crisis is not handled properly. Due to negative emotions, stakeholders can decide to break off interactions with an organization or engage in negative word of mouth about the organization.

Tweets provide useful real-time information for decision-making and communication during crises [11], [12]. However, given the vast amount of data online, this information cannot be directly used. Applying sentiment analysis is one option to make this vast amount of information manageable and usable. By using sentiment analysis, tweets expressing positive and negative emotions can be detected and analyzed against each other. Contrary to sentiment analysis, which classifies tweets as positive or negative, affect analysis or emotion recognition classifies tweets as belonging to a specific emotional state (e.g., happiness, anger) [13]. Since it is a multinomial classification problem rather than a binary classification problem, affect analysis is even more challenging than sentiment analysis [14]. Most systems for automatic analysis of emotions are based on the six basic emotions of Ekman [15], namely *anger, fear, sadness, enjoyment, disgust,* and *surprise*. Strapparava and Mihalcea [13] constructed a large data set of news headlines that were annotated with these basic emotions and developed a binary classifier for each emotion. Their experiments show that the classification performance varies strongly between the different emotion categories (F= 4.68 for *disgust* vs. F= 32.78 for *joy*). However, the Ekman scale does not account for the typical emotions expressed in organizational crises. In order to account for these crisis-related emotions, Jin et al. [9] proposed an emotion framework in which they identified three clusters of crisis emotions: i) attribution-independent emotions, which consist of anxiety, fear, apprehension, and sympathy; ii) external-attribution-dependent emotions, including disgust, contempt, anger, and sadness; and iii) internal-attribution-dependent emotions, which consist of embarrassment, guilt, and shame. Attribution-independent emotions are emotions people feel toward a crisis situation; external-attribution-dependent emotions are emotions people feel about an organization in a crisis; and internal-attribution-dependent emotions are emotions people feel for themselves as stakeholders involved in a crisis.

III. Automatic Classification of Sentiment and Emotion in Crisis-related Microposts

In this section, we report on the data and different experiments we performed on the tweets related to the crash of a Germanwings airplane in the French Alpes.

A. Dataset

On Tuesday, March 24, 2015, around 10:41 Central European Time, an Airbus A320-200 crashed in the French Alps, 100 kilometres northwest of Nice. It concerned Flight 9525, an international passenger flight from Barcelona-El Prat Airport in Spain to Düsseldorf Airport in Germany. The flight was operated by Germanwings, a low-cost airline owned by Lufthansa. First, the crash was assumed to be an accident. On March, 26, however, the French Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile discovered after analyzing the aircraft's flight data recorder that co-pilot Andreas Lubitz deliberately crashed the aircraft. Two pilots, four cabin crew members, and 144 passengers were on board of the aircraft. No one survived the crash. In the week after the crash, evidence was found that Lubitz suffered from a psychosomatic illness and that he was taking prescription drugs.

For this paper's study, a corpus of English tweets was collected. The Twitter search facility was used in order to find all English posts, made by any Twitter user, that contained the hashtag '#GermanWingsCrash'. Given the vast amount of tweets, a random selection was made of a maximum of 25 tweets per hour, posted between March 24, 2015 and April 6, 2015. A total of 5,490 English tweets were harvested.

 ISBN: 978-1-61208-519-7

B. Sentiment Classification

In order to determine the sentiment that was conveyed in the tweets, we used a machine learning approach to sentiment detection [16] to determine the polarity of the tweets. We used the system developed by Van Hee et al. [16] in the framework of the SemEval-2014 Task 9 on sentiment analysis in Twitter. First, linguistic preprocessing (including tokenization, PoS-tagging, lemmatization and dependency parsing) was performed on the datasets. Then, a number of lexical and syntactic features were implemented: n-gram features, word shape features (e.g., the number of capitalized words), lexicon features, syntactic features (e.g., Part-of-Speech information), named entity features and PMI features (PMI values indicate the association of a word with positive and negative sentiment). After performing feature selection experiments, it was discovered by Van Hee et al. [16] that features based on n-grams, sentiment lexicons, and Part-of-Speech tags were most contributive for labelling a message or an instance of that message as positive, negative, or neutral.

The system labelled 676 tweets as positive, 2,815 tweets as negative, and 1,999 tweets as neutral. Given that this corpus contains tweets referring to the crash, the large number of negative tweets is not surprising. In order to assess the quality of the automatic labelling, we manually annotated a corpus of 200 tweets with polarity information (see Table II). We observed a classification accuracy of 73.17% for the negative class, 26.92% for the positive class and 64.83% for the neutral class. The total system accuracy amounted to 63.32%. It can be concluded that the system particularly made mistakes with regard to the positive class label. This could be explained by the fact that the system has been trained with Twitter messages on a variety of general topics and not with crisis-related tweets. As a result, the training datasets delivered in the framework of the SemEval-2014 shared task contained more positive tweets (38.20%). Moreover, it can be concluded that the system performed best with regard to the negative class label. This is a significant advantage in crisis situations, in which the detection of negative emotions is highly important.

TABLE II. Polarity detection classification accuracy

Polarity	*# tweets*	*Accuracy*
Positive	26	26.92%
Negative	82	73.17%
Neutral	92	64.83%

C. Towards Emotion Detection

Research on understanding emotions in crisis-related tweets and more specifically to pinpoint those tweets which might cause organizational harm, is scarce. Consequently, no system was available yet to detect crisis-related emotional content in tweets. In order to better understand the types of emotions expressed, we took the 200 tweets which were manually labeled with polarity information and also labeled them with emotions. For this purpose, the scale of Jin et al. [9] was used, since it was specifically developed for measuring the publics' emotions in organizational crises. This crisis emotion scale consists of thirteen discrete emotions, being *anger, anxiety, apprehension, confusion, contempt, disgust, embarrassment, fear, guilt, sadness, shame, surprise,* and *sympathy*. For the annotation, we grouped a number of emotions as they were difficult to differentiate, namely anxiety and fear, and shame and embarrassment. Tweets that conveyed an emotion that did not occur in Jin et al.'s crisis emotion scale were labelled as *other*.

For the tweets expressing a positive or a negative sentiment, we tagged the emotional content as one of the classes anger, fear, apprehension, confusion, contempt, disgust, embarrassment, guilt, sadness, surprise, sympathy and other. Table I gives an overview of the occurrence of these emotion classes in our English gold standard corpus. Sympathy, anger and contempt are the emotions that were most frequently expressed in the data. No tweets conveying embarrassment or guilt were found in our gold standard corpus. For each emotion class, an example tweet is represented in the last column.

IV. Reflections

Important to note is that the current annotation scheme fails to detect the object of the expressed emotion. While the sympathy emotions are mostly expressed towards the family members of the victims, the tweets expressing anger and contempt have a completely different object, most often the co-pilot that deliberately crashed the plane (e.g., "F'ing lunatic. Kill yourself, not a load of passengers! #flight9525"). In order to make emotion detection really viable for business intelligence, a more fine-grained approach in the annotation of the external-attribution-dependent emotions should be taken into account.This way, not only emotions are labelled, but also the objects of these emotions (and maybe also the senders of these emotions), a tendency we also observe in the domain of sentiment analysis (see Pontiki et al. [17]). Important to know for crisis managers of companies such as Germanwings and Lufthansa is how people report on their organizations, something which is not being covered by the current annotation scheme.

A shallow analysis of the tweets reveals that many of them refer to the aviation sector as a whole ("sad day in aviation again", "another flight crash", "far too many planes going down", "In The Wake Of The #GermanWingsCrash Crash, Should You Trust Low-Cost Airlines?", etc.), which has a general image problem. Two main criticisms were specifically targeted towards both Germanwings and Lufthansa: i) that they did not immediately release the names of the pilots (which was done a day after the crash) and ii) that the cockpit should always have two persons present (this two-in-the-cockpit rule was very soon adopted). If this criticism would have been automatically detected, then crisis managers would have had a guiding tool for adequate crisis communications while the crisis was unfolding. This is how accurate emotion detection in the future could make a difference, ultimately reducing reputation harm for organizations.

V. Conclusion

The main goal of this study was to investigate the extent to which automatic sentiment analysis techniques can be used to detect crisis emotions on Twitter. We conclude that the sentiment analysis system performed better on negative tweets when compared to tweets expressing a positive emotion. Although during crises negative emotions are most prevalent and relevant for crisis emotions to focus upon, positive emotions should not be neglected. To have a better understanding of how stakeholders respond to crisis victims (e.g., by showing sympathy), to the organization itself –both at the beginning of a crisis, while the crisis is unfolding and after crisis communication has been made (e.g., apologies, condolences), it is also crucial to have a more fine-grained classification of specific crisis-related emotions. In order to allow for the future development of such automatic procedures, we conducted a small corpus analysis for which we manually labeled our corpus with crisis-related emotions. We found that sympathy and anger were the most frequently expressed emotions in the English gold standard corpus in the case of the Germanwings crash. We also observed that the annotation of crisis-related emotions in the tweets was insufficient to support organizational crisis communication. To further enhance the usefulness of automatic (crisis) emotion detection on social media, future studies should work on the classification of contextual information, such as the object and characteristics of the sender of the crisis emotion.

REFERENCES

[1] Y. Jin, B. F. Liu, and L. L. Austin, "Examining the role of social media in effective crisis management: the effects of crisis origin, information form, and source on publics' crisis responses," Communication Research, vol. 41, no. 1, 2011, pp. 74–94.

[2] A. Schwarz, "How publics use social media to respond to blame games in crisis communication: The Love Parade tragedy in Duisburg 2010," Public Relations Review, vol. 38, no. 3, 2012, pp. 430–437.

[3] S. R. Veil, T. Buehner, and M. J. Palenchar, "A work-in-process literature review: incorporating social media in risk and crisis communication," Journal of Contingencies and Crisis Management, vol. 19, no. 2, 2011, pp. 110–122.

[4] T. Coombs, Ongoing crisis communication : planning, managing, and responding (3 ed.). Thousand Oaks, Calif.: SAGE, 2012.

[5] J. Bryson, "What to do when stakeholders matter: Stakeholder identification analysis techniques," Public Management Review, vol. 6, no. 1, 2004, pp. 21–53.

[6] T. Coombs, "Impact of past crises on current crisis communication: Insights from Situational Crisis Communication Theory." Journal of Business Communication, vol. 41, no. 3, 2004, pp. 265–289.

[7] ——, "Protecting organization reputations during a crisis: The development and application of Situational Crisis Communication Theory," Corporate Reputation Review, vol. 10, no. 3, 2007, pp. 163–176.

[8] Y. Jin, "Making sense sensibly in crisis communication: How publics' crisis appraisal influence their negative emotions, coping strategy preferences and crisis response acceptance," Communication Research, vol. 37, no. 4, 2010, pp. 522–552.

[9] Y. Jin, B. F. Liu, D. Anagondahalli, and L. Austin, "Scale development for measuring publics' emotions in organizational crises," Public Relations Review, vol. 40, 2014, pp. 509–518.

[10] H. Kim and G. Cameron, "Emotions matter in crisis: The role of anger and sadness in the publics' response to crisis news framing and corporate crisis response," Communication Research, vol. 38, no. 6, 2011, pp. 826–855.

[11] T. Heverin and L. Zach, "Microblogging for crisis communication: examination of twitter use in response to a 2009 violent crisis in the seattle-tacoma, washington area," in Proceedings of the 7th International ISCRAM Conference, Seattle, USA, 2010, pp. 1–5.

[12] N. D. Sreenivasan, C. S. Lee, and D. H.-L. Goh, "Tweet me home: Exploring information use on twitter in crisis situations," in Proceedings of the 14th International Conference on Human-Computer Interaction, Orlando, Florida, USA, 2011, pp. 1–10.

[13] C. Strapparava and R. Mihalcea, "Learning to identify emotions in text," in Proceedings of the 2008 ACM symposium on Applied computing, 2008, pp. 1556–1560.

[14] F. Brynielsson, J.and Johansson and A. Westling, "Learning to classify emotional content in crisis-related tweets," in Proceedings of the 11th IEEE International Conference on Intelligence and Security Informatics (IEEE ISI), Seattle, Washington, USA, 2013, pp. 33–38.

[15] P. Ekman, "An argument for basic emotions," Cognition & Emotion, vol. 6, no. 3-4, 1992, pp. 169–200.

[16] C. Van Hee, M. Van de Kauter, O. De Clercq, E. Lefever, and V. Hoste, "LT3: Sentiment classification in user-generated content using a rich feature set," in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 2014, pp. 406–410.

[17] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. AL-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, V. Hoste, M. Apidianaki, X. Tannier, N. Loukachevitch, E. Kotelnikov, N. Bel, S. M. Jimenez-Zafra, , and G. Eryigit, "Semeval-2016 task 5: Aspect based sentiment analysis," in Proceedings of the 10th International Workshop on Semantic Evaluation, San Diego, California, USA, 2016, pp. 19–30.

 ISBN: 978-1-61208-519-7