Automatic Emotion Detection in Social Media for on the fly Organizational Crisis Communication

Karolien Poels
Department of Communication Studies
University of Antwerp
Antwerp, Belgium
email: karolien.poels@uantwerpen.be

Veronique Hoste
LT3 Language and Translation Technology Team
Ghent University
Ghent, Belgium
email: veronique.hoste@ugent.be

Abstract—**We present early stage research ideas on automatic emotion detection on social media (Twitter) during an organizational crisis and how this can be used for rapid and effective organizational crisis communication.**

Keywords—emotion detection; social media; natural language processing; organisational crisis; crisis communication.

I. INTRODUCTION

Many profit and non-profit organizations use social media like Facebook and Twitter for two-way communication with their stakeholders [1]. It is obvious that social media play an increasing role in organizational crisis communication [2] [3]. A crisis is "the perception of an unpredictable event that threatens important expectancies of stakeholders and can seriously impact an organization's performance and generate negative outcomes" [3]. Examples are product failures, or unethical practices. Effective crisis communication includes identifying stakeholders (i.e., a person or a group that is affected by or can affect an organization – e.g., customers, shareholders, journalists, the general public) and entering in discussion with them [3]. Social media present unique opportunities for crisis communication, such as the absence of journalistic gatekeepers, by which an organization can communicate its own, unbiased account [4]. Also, social media enable informing stakeholders more frequently and faster than traditional media. At the same time, stakeholders respond more directly, open and rapidly and do also expect an evenly fast response from organizations. It is thus utmost important that organizations know how their different stakeholders are feeling when a crisis is unfolding in order to respond efficiently and avoid (further) reputational harm.

In times of crises, stakeholders typically experience various emotions (e.g., anger, fear, worry, relief). Jin et al. [5] have recently classified publics' emotions in response to organizational crises depending on how (much) crisis responsibility is attributed to the organization. An important challenge is that, given the large share of comments posted by stakeholders during a crisis, it is impossible to manually monitor all these posts. A fine-grained automated detection method to tap into emotions communicated through social media on the fly is needed, but currently still underdeveloped.

Our goal is to study how emotions are experienced and expressed through social media in times of crises and how these emotions can be detected automatically with the ultimate aim to adapt corporate crisis response strategies in a rapid and efficient way. This research endeavor requires an interdisciplinary approach. There will be close cooperation between two disciplines: communication sciences (CS), integrating emotion theory and social media dynamics in crisis communication models, and natural language processing (NLP) for text mining of emotions in social media communication. We focus mainly on Twitter as one of the most important social media, especially for crisis communication [2].

In the Section 2 of this ideas paper, we discuss the two core objectives that need to be tackled to realize the future development of automatic emotion detection in organizational crises on social media.

II. RESEARCH OBJECTIVES

A. Objective 1: The Nature of Stakeholder Emotions in Social Media Crisis Communication

There is consensus that emotions expressed by the stakeholders during crises are a crucial component to consider in crisis communication [5]. However, there is reason to assume that some emotions will be more easily expressed through social media than others. Apart from some fragmented studies, a solid integration of emotion theory in social media crisis communication models is still lacking.

Several new steps need to be taken. Emotions are complex and multifaceted phenomena. Therefore, when studying how emotions are expressed through social media in crisis situations, it is crucial to first have an accurate view on *what* emotions are, *why* they are triggered in response to organizational crises, and subsequently *why, when* and *how* they are most likely to be expressed through social media. ***This critical evaluation of emotion theory and the dynamics***

of social media communication and Twitter more concretely is needed to build a conceptual model integrating emotions in social media crisis communication.

B. Objective 2: Automatic Detection of Crisis Related Emotions in Tweets

Until now, most work on subjectivity analysis in the domain of NLP has concentrated on discovering whether a review, tweet, sentence, or specific "object" (person, product, organization, etc.) is regarded in a positive or negative manner by a specific "source". This task has been given many names, from opinion mining, to sentiment analysis, review mining, attitude analysis, appraisal extraction, etc., and exists in different granularities ranging from coarse-grained sentiment analysis at the text level to fine-grained aspect-based sentiment analysis. From a methodological point of view, it is generally assumed that the semantic properties of words are good predictors of the semantic characteristics of the phrase or text that contain them. As a result, a lot of effort has been invested in the manual and (semi-)automatic development of lists of words indicative of sentiment [6] [7], and in the development of systems which both exploit these external lexicons and the lexical information present in the data (e.g., unigrams, bigrams, etc.) to determine for a given text, sentence or phrase whether it is positive, negative or neutral see [8] [9] for an overview. A similar methodology relying on external lexicons, e.g. [10], and automatically derived lexical knowledge has also been used in the more fine-grained approaches aiming at the automatic detection in text of emotions, in most cases restricted to a discrete set of emotions such as the 6 "universal" emotions distinguished by Ekman [11]: anger, disgust, fear, sadness, joy and surprise [12] [13] or the 12 emotions selected in the context of a shared task on emotion classification in suicide notes [14]. The basic Ekman emotions have also recently been investigated in tweets related to natural disasters such as Hurricanes [15].

Fine-grained emotion detection, however, remains a largely understudied research area in the NLP domain. Furthermore, there are no approaches modeling emotions related to organizational crisis in tweets more specifically. *In order to move beyond superficial emotion modeling and to find a machine learnable operationalization of the types of emotions that are most diagnostic for an organizational crisis, an interdisciplinary approach is required.*

III. RESEARCH CHALLENGES

The integration of emotion theory in social media crisis communication models is crucial to the development of a supervised machine-learning model that can detect crisis-related emotions in tweets. Key to the success of these models is 1) to determine relevant crisis emotions and how they are expressed in social media, 2) to define the optimal set of text features (e.g., detecting words indicative of a given emotion,

modeling modality, modeling sarcasm, etc.) for the detection of a given emotion. Important research questions to be investigated are: How well defined are the different emotions or do they represent a continuum? How learnable are the annotated emotions? How to include time series information? Should we weigh twitter users differently depending on their role as a stakeholder (e.g., affected customers, journalists, random twitter users)? How should we handle retweets? How generic are the learned models?, etc.

REFERENCES

[1] J. N. Sutton, "Social media monitoring and the democratic national convention: New tasks and emergent processes," J. of Hom. Sec. and Em. Man., 6(1), 2009, pp. 1-20.

[2] F. Schultz, S. Utz, and A. Göritz, "Is the medium the message? Perceptions of and reactions to crisis communication via twitter, blogs and traditional media," Public Rel. Rev., 37, 2011, pp. 20-27.

[3] T. W. Coombs, "Ongoing crisis communication: planning, managing and responding (4th edition)," Thousand Oaks, California: Sage publications, 2015.

[4] 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 public's crisis responses," Comm. Res., 2011, pp. 1-21.

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

[6] J. Wiebe, M. Bruce, F. Rebecca, and T. O'Hara, "Development and use of a gold standard data set for subjectivity classifications," Proc. of ACL-99, 1999, pp. 246-253.

[7] V. Jijkoun and K. Hofmann, "Generating a non-English subjectivity lexicon: relations that matter," Proc. of EACL-2009, 2009, pp. 398-405.

[8] A. Balahur, "Methods and Resources for Sentiment Analysis in Multilingual Documents of Different Text Types", PhD thesis, Univ. of Alicante, unpublished, 2011.

[9] B. Liu, "Sentiment Analysis and Opinion Mining," Morgan & Claypool Publishers, 2012.

[10] J. Staiano and M. Guerini, "DepecheMood: a Lexicon for Emotion Analysis from Crowd-Annotated News," Proc. of ACL-2014, 2014, pp. 427-433.

[11] P. Ekman, "An argument for basic emotions," Cognition and emotion, 6(3/4), 1992, pp.169-200.

[12] C. Strapparava and R. Mihalcea, SemEval-2007 Task 14: Affective Text. Proc. of SemEval-2007, 2007, pp.70–74.

[13] J. Bellegarda, "Emotion Analysis Uting Latent Affective Folding and Embedding," Proc. of the NAACL HLT 2010 Worksh. on Computational Approaches to Analysis and Generation of Emotion in Text, 2007, pp. 1-9.

[14] B. Desmet and V. Hoste, "Emotion Detection in Suicide Notes," Expert Systems with Applications, 40 (16), 2004, pp. 6351-6358.

[15] J. Brynielsson, F. Johansson, and A. Westling, A, "Learning to Classify Emotional Content in Crisis-Related Tweets," ISI 2013.