International Journal of Information Systems for Crisis Response and Management
Volume 8 • Issue 3 • July-September 2016

Improving the Utility of Social Media Data to Emergency Responders through Emotional Content Detection

Shane Halse, Pennsylvania State University, University Park, PA, USA
Andrea H Tapia, Pennsylvania State University, University Park, PA, USA

ABSTRACT

In the following paper, we will present an alternate method for the detection of emotional content within social media data. Current research has presented the traditional bag-of-words method in which a predefined corpus is used to measure the emotional context of each word within a message. Here we present a method in which a small subset of the data is labeled to generate a corpus which is then used to detect emotional content within the data. This research is being conducted on the dataset from hurricane Sandy in 2012. Our findings show an improvement upon the bag-of-words method. These findings would further the current research in improving the utilization of social media data within crisis response. In doing this we allow the average citizen to provide beneficial data to those in decision making roles.

KEYWORDS

Emotion, Hurricane, Sandy, Sentiment, Trust, Twitter, Usefulness

INTRODUCTION

We as humans are social creatures, and this sociability breeds empathy. However, machines lack the ability to empathize with us. It is this lack of empathy that makes them both ideal for big data analysis and yet still limits them when attempting to examine human behaviors. One such behavior which we will examine in this paper is that of cognitive dissonance between machine and human assessment. Imagine a situation in which a large-scale disaster has occurred, people are left homeless, out in the elements, and scared. Within this situation, you receive three messages; "help me!", "help me I am so worried", and "I need help because of this stupid storm". Even with no further information we are often inclined to empathize with those who are scared. Thus, even though the three messages are a request for a help, we are more likely to trust the one who has exhibited vulnerability (McEvily et. al.2003). However, for various reasons, a machine assessment of these same three messages may speak to lower level of trustworthiness of the emotional message, potentially causing conflict between human and machine estimates, thus leading to the cognitive dissonance effect.

There are numerous barriers that need to be overcome with the adoption and use of social media data during a crisis. However, even with these barriers, this data has been used during crisis response situations (Tapia, Bajpai, Jansen, Yen, & Giles, 2011) (Hughes & Palen, 2009). One downside to using data found in social media is the amount of data that needs to be analyzed during a crisis event. For example, if we take the social media platform Twitter, during the Sandy hurricane (Oct 27, 2012 – Nov 1, 2012) over 20 million tweets with the terms "sandy" and "hurricane" were posted (Twitter 2012). We can see from the vast amount of data generated during the 6 days during the hurricane that

DOI: 10.4018/IJISCRAM.2016070102

having a team manually analyze the data would be too time consuming and inefficient. Another issue with utilizing the information found within social media are the issues with credibility, liability, and organizational process and procedure (Tapia et al. 2011; Tapia, Moore, and Johnson 2013; Starbird and Palen 2013). One potential solution has been to investigate the use of machine learning in order to identify those tweets deemed the most trustworthy, thus taking the first steps to prioritize the data to be vetted (Mendoza, Poblete, & Castillo, 2010; Thomson et al., 2012).

Many features of social media data have been found to contribute to trust such as; message meta data, message content, author, network and links, however they are spread across multiple domains (Morris, Counts, Roseway et al., 2012; Gupta, Kumaraguru, & Castillo, 2014). One of this paper's goals is to enhance features and methods found in past research through the detection of the emotional content within the messages to further improve a machines trust assessment. It is these emotions that play a critical role when a human is required to "believe" the machine's assessment. Ironically, much research has been done in terms of detecting emotional sentiment within a message through the use of machine learning (Pak & Paroubek, 2010; Alm, Roth, & Sproat, 2005; Halse, Tapia, Squicciarini et al., 2016). However, these techniques have not yet been utilized within the domain of trust, especially within crisis management. It is in combining these trust assessment features combined with the emotional sentiment of a message that we hope to improve the quality of the messages provided to the crisis responder.

Currently the automation of social media trust assessment, while continuously improving, has not yet led to its adoption at a large scale (Tapia & Moore, 2014). It is the goal of this paper to investigate the use of machine learning to evaluate the emotional content of a message, and how the addition of this evaluation can improve a machines trust assessment. In addition, by using machine learning to evaluate the emotional content of a message, we can investigate the cognitive dissonance formed between a machine generated trust assessment and a human's perception of trustworthiness. That is, it will provide the opportunity to investigate and explain why some messages deemed untrustworthy by one source such as a machine's trust assessment and yet may still be perceived as trustworthy by the human.

We know that emotion has an effect on the perceived trustworthiness of a message and that there are currently ways in which trustworthiness can be assessed through non-message content related features. Thus, the idea behind this research would be to explore this further. That is, we want to start by investigating the idea of how emotional key words can change the tone/interpretation of messages. Next, we want to develop or find ways in which we can utilize existing tools (such as sentiment analysis) to discover the underlying tone of a message in order to improve a machines representation of it. This would allow us to determine which message may be more trustworthy than others and provide a basis upon which we can mark each of these messages for further vetting. In addition, by discovering the emotional tone within each message we'd be able to aggregate this data to provide an overall emotional status of the community affect by the disaster. This emotional status would serve as a guide to human decision makers.

Prior Research and Background

In a time of crisis, there are many decisions that need to be made. These decisions can range in importance from what type of coffee to order to decisions that may costs lives. Unfortunately, information used while making these decisions may not always be reliable or may conflict with information from other sources. It is up to the decision maker to decide which path he or she will take and which information to base these decisions on. Because of this we become susceptible to bad decision making especially when the information available to us has an emotional content or

Related Content

Comparing Different Crowd Emergency Evacuation Models Based on Human Centered Sensing Criteria

Jaziar Radianti, Ole-Christoffer Granmo, Noureddine Bouhmala, Parvaneh Sarshar and Jose J. Gonzalez (2014). *International Journal of Information Systems for Crisis Response and Management (pp. 53-70).*

www.igi-global.com/article/comparing-different-crowd-emergency-evacuation-models-based-on-human-centered-sensing-criteria/128221?camid=4v1a

Use Team Building to Make the Most of Your Public-Private Partnerships

Martin Negron (2015). *Emergency Management and Disaster Response Utilizing Public-Private Partnerships (pp. 121-135).*

www.igi-global.com/chapter/use-team-building-to-make-the-most-of-your-public-private-partnerships/124654?camid=4v1a