Emotion analysis of reaction to Terrorism on Twitter

Jonathas G. D. Harb[1], Karin Becker[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

jonathasgabriel05@gmail.com, karin.becker@inf.ufrgs.br

Abstract. *Terrorism events impact people in several manners. Reactions may include losing sense of safety and experiencing angry and fear, among others. The social media has become an important mean where people express themselves. We target Twitter to investigate the emotional reaction people have to terrorism events. For this purpose, we analyze emotions in tweets along with demographic data. Tracking emotional reaction can help in defining specific assistance programs. In our approach we collect a corpus of tweets related to two terrorism events, classify emotions, extract user location and estimate user age and gender with use of available tools. Results showed an emotion shift due to the events and a difference on the reaction from one event to another.*

1. Introduction

Terrorism in all its forms remains a constant threat that continues to be present in the global agenda and raises questions concerning prevention and consequences. In general, terrorism involves the use, or threat of use, of violence as an attempt to achieve some social or political effect. The goal of terrorism is to create instability by propagating fear, arousal and uncertainty on a wider scale than those achieved by targeting a single victim (Horgan 2014). Terrorism attempts are becoming more frequent and diverse, and reactions to terrorism events include, among others, loosing sense of safety, feeling helpless, experiencing anger and fear. Tracking user emotions can help authorities to define and provide specific assistance programs for coping with it.

Today's widely accessible social micro-blogging platforms such as Twitter are increasingly being used on global scale to publish content and express emotions and opinions on a daily basis. This large volume of information are being explored by data science area for several purposes, such as to identify sentiment and emotions expressed in tweets (Mohammad et al. 2015; Mohammad 2012), to monitor how people feel about specific topics (Wan and Paris 2015), to predict information flow size and survival following specific events (Burnap et al. 2014), to analyze social connections (Lerman et al. 2016), and to study engagement to context-specific tweets (Suttles and Ide 2013), among others.

One opportunity is to explore such information to investigate users' emotional reaction to terrorism events, through emotion analysis. Sentiment Analysis is the field of study that analyzes people's opinions, appraisals, evaluations, attitudes, and emotions from written language (Liu 2012). Emotion mining involves identifying emotion bearing words/expressions in texts and classifying them according to an emotion model (Munezero et al. 2014). Common approaches for sentiment classification are the adoption of emotion lexicons and supervised learning over emotion-labeled data resources. These resources are less abundant, when compared to polarity classification. Popular emotion models are basic emotions and VAD (valence, arousal and dominance) (Munezero et al. 2014).

Sentiment analysis in tweets is difficult due to its unstructured and informal short text. By design, users have a limited number of 140 characters, and typically tweets contain casual text with errors (spelling, grammar, etc), abbreviations, internet-specific terms, etc. In addition, working with emotions in tweets is a much less studied problem by the literature due to the lack of labelled data (Hasan et al. 2014).

Previous works have focused on sentiment analysis in texts, such as tweets, with different goals. While authors in (Mohammad et al. 2015; Mohammad 2012) applied different labelling techniques, others have proposed novel approaches for classifying tweets into sentiment categories (Wang et al. 2012; Purver and Battersby 2012; Bravo-Marquez et al. 2016; Kim 2014). Works have also focused on analyzing social aspects of emotion in twitter (Kim et al. 2011) and on using demographic data to characterize social relations (Lerman et al. 2016) and population mobility patterns (Gallegos et al. 2015). Emotional reaction to specific topics or events was the focus in (Wan and Paris 2015). However, studies on sentiment analysis to investigate emotional reaction to terrorism in Twitter are still lacking.

In this paper, we aim to investigate and help understand the emotions people express about terrorism events with help of demographics data. For that purpose, we collected and analyzed data on two terrorism events that occurred in England, in order to answer the following questions:

- Q1: Is there an emotion shift due to terrorism events?
- Q2: Do different terrorism events raise the same emotional reaction?
- Q3: Do user location have an impact on the emotional reaction?

To answer these questions, we collected a corpus of terrorism-related tweets, identified the presence of emotions with deep learning methods, and determined demographic information of the respective users, such as location, age and gender. To the best of our knowledge, this is the first work that investigates emotional reaction in Twitter in the specific context of terrorism.

The remaining of this paper is structured as follows. Section 2 describes related work. Section 3 describes the methods and materials used for providing answers to our research questions. Section 4 describes our experiments to find a suitable model for emotion prediction. Section 5 presents the analysis performed over the data to answer our questions. Finally section 6 presents the conclusion and opportunities for future works.

2. Related Work

Sentiment analysis was target of several works for different purposes. A series of approaches for sentiment classification and data labelling were presented in (Mitchell et al. 2013; Anderson 2005; De Choudhury et al. 2016; Lotan et al. 2011). The work presented in (ElSherief et al. 2017) applied distant supervision and machine learning methods such as Naive Bayes (NB) and Maximum Entropy (ME) for sentiment classification. Kim (Kim 2014) applied deep learning in natural language text by training a convolutional neural networks (CNN) on top of pre trained word embeddings. Models were evaluated against several datasets and the results outperformed several state of the art methods in the majority of the experiments.

Previous works have also explored the social questions involving tweet sentiment analysis. Lerman et al. (2016) analyzed a large corpus of georeferenced tweets in order to study the structure of social connections people form online. Their work collected tweets from US areas, linked these tweet locations to corresponding Census data and estimated tweet sentiment into negative and positive through SentiStrenght[1]. In their analysis, they were able to identify groups that expressed more positive emotions as well as groups where negative emotions were predominant. The structure of social connections of these groups as well as demographic data helped in explaining such findings.

Suttles et al. (2013) studied user engagement with Gender-based violence (GBV) related posts in Twitter. Age, gender and linguistic attributes, including emotions, were analyzed. Their work reported how users engage with GBV tweets based on favoriting and retwitting metrics. Descriptive statistical analysis was applied to identify age and gender of tweeters. Tweet sentiment was extracted from the LIWC software[2]. They found that users engage more to GBV related tweets than to generic tweets and that the engagement is not uniform across genders and ages. Moreover, anger was often predominant in the GVB content context.

Gallegos et al. (2015) used data from Foursquare service to identify US metropolitan areas that people use to check-in. These areas were then analyzed with help of demographics data. Their results revealed that areas with many check-ins have happier tweets and therefore encourage other people to connect to these places. They reported that such results provided more information on human mobility patterns.

Despite all cited contributions, just a few studies have focused on emotion mining in their tasks and none of the them have studied sentiment analysis in a specific context as we do by targeting terrorism events.

3. Materials and Methods

3.1. Dataset

We decided to target two terrorism events that occurred in the United Kingdom. This choice was motivated by two factors. First, we focused on the English language, in order to benefit from many tools and functions available for natural language processing. Second, to study emotional reaction for different events in a same region, as England was the target of a few attacks in 2017.

The first event was the Manchester Arena bombing[3], which took place on May 22th, 2017 in Manchester, when people were leaving a concert of Ariana Grande. The second one was the London Bridge attack, occurred in London on June 3rd 2017, where a van left the road and struck a number of passing by pedestrians[4].

Data collection must involve tweets from the past, as the occurrence of a terrorism event is unpredictable. As the Twitter official streaming API does not allow to collect

[1]http://sentistrength.wlv.ac.uk/
[2]http://liwc.wpengine.com
[3]https://en.wikipedia.org/wiki/Manchester_Arena_bombing
[4]https://en.wikipedia.org/wiki/June_2017_London_Bridge_attack

Table 1. Query Terms, Dates and Dataset per Event

Event Name	Query terms	Period	BEFORE (#tweets)	AFTER (#tweets)
#prayformanchester	#prayformanchester, ”Manchester”	05-20-2017 to 05-24-2017	BM (5,351)	AM (25,010)
#londonbridge	#londonBridge, ”London”	06-01-2017 to 06-05-2017	BL (20,379)	AL (29,656)

tweets from the past, we used an open source project[5] written in Python, which bypasses some of the limitations of Twitter API. As parameters, we set query search terms combined with boundary dates.

For each targeted event, we collected tweets two days before the event, the actual day it happened, and two days after the event. In this way, we were able to analyze not only emotion reaction, but also a possible emotion shift due to the events. To define search terms to collect tweets, we analyzed raw data gathered from the web, trending topics, as well as samples extracted using the official Twitter API on the respective dates. We found recurrent hashtags for each one of the events, namely #prayformanchester for the Manchester attack and #londonbridge for the London one. We assumed these hashtags were representative due to their major predominance in tweets referring to these events (nearly 10 times more frequent). Tweets collected two days before the events were queried by the keywords ”Manchester” and ”London”. We considered these tweets as representative, as we observed that these keywords were commonly used to tweet about citizen's thoughts on diverse topics such as football teams, universities, and daily news regarding these cities, among others. Table 1 shows queries search terms and boundary dates for tweet collection.

Data pre-processing involved traditional steps, such as the removal of hyperlinks, hashtags (because they did not provide useful information other than identifying the events), mentions to other users, special marks and symbols (&, /, ¿ _, etc). In addition, we applied an English dictionary[6] to filter out tweets with too many misspelled words and non English ones. These actions resulted in four datasets (shown in table 1): BM (before Manchester) containing 5,351 tweets, BL (before London) containing 20,379 tweets, AM (after Manchester) containing 25,010 tweets, and AL (after London) containing 29,656 tweets. The structure of these datasets is identical and include, among others, the filtered tweet text and the tweet ID.

We characterized the demographics of the data in terms of location, gender and age. For location, we observed that less than 1% of the collected tweets were georeferenced. Thus, we assumed that the location of the tweet would be extracted from the users' profile as in (Sakaki et al. 2010). Our original idea of analyzing sentiment per city could not be accomplished due to the low number of tweets for comparison. On the other hand, in analyzing location by countries and larger regions, we found out that a representative number of locations from the UK and the US were present in the dataset Therefore, we filtered locations in three categories: locations from the UK, locations from the US and ”other locations”. As the location in each user profile is a simple text without any validation, we compared each declared location against a list of cities from the UK and the US to include in these two categories. As in (ElSherief et al. 2017), gender and age were esti-

[5]https://github.com/Jefferson-Henrique/GetOldTweets-python
[6]https://github.com/dwyl/english-words

Table 2. Gold Standard: Number of labelled tweets per category

Emotion	Anger	Disgust	Fear	Sadness	Surprise	None
# tweets	82	116	85	179	71	74

mated using Face++[7]. Face++ provides an API that analyzes face related attributes based on machine learning, and experiments evaluated an accuracy of 85% (Fan et al. 2014)

3.2. Gold Standard

Our work focuses on five out of the the six basic emotion categories defined by Eckman (Ekman and Friesen 1982). We focused on negative emotions only, because we assume people are not likely to express positive emotions (such as happiness) in reaction to terrorism events[8]. The emotion categories considered therefore include anger, fear, sadness, surprise and disgust. Our approach considers that a given tweet is included in one and only one of the emotion categories.

To train a model for emotion prediction, an emotion labeled dataset is required. As domain-related datasets tend to provide the best results (Liu 2012), we created a specific terrorism gold standard for the task. Tweets were labeled according to each emotion category considered, plus an extra ”none” category. This was accomplished using Amazon Mechanical Turk[9].

First, one of the authors annotated 967 tweets with the considered 5 emotion labels, based on the presence of emotion keywords and expressions. For example, the tweet ”*Deeply saddened by the loss of 22 beautiful lives. we should not live like this. They did not deserve to die*” was labelled as sadness due to the expression ”deeply saddened”; the tweet ”*It’s so scary to not feel safe in this World*” was labelled as fear due to the expression ”It’s so scary”, and so on. We started from a randomly selected set of tweets, discarding the ones that did not contain an unquestionable emotion word/expressions, and labeling otherwise. This task was performed until we reached a minimum of 100 tweets per emotion. This procedure resulted in relatively well balanced sets. Afterwards, we created a HIT (Human Intelligence Task) with these tweets, where annotators were asked to determine which emotion best described a tweet, given a set of categories as options (anger, fear, disgust, sadness, surprise, none). We instructed annotators to choose the primary emotion if more than one emotion could be identified, and to choose ”none” if no emotion could be clearly determined. We targeted the HIT to two master annotators, so that we would have three annotators in total, considering one of the authors. According to Amazon, master annotators typically have a 90% or more of accuracy rate. We filtered out tweets in which there was a disagreement between all the three annotators, and retained those with at least two agreements. The results, composed of 607 tweets, are displayed in Table 2, which we consider as our ground truth for validating the emotion prediction model.

[7]https://www.faceplusplus.com/
[8]https://www.paulekman.com/blog/our-emotional-reactions-terrorism/
[9]https://www.mturk.com/

3.3. Classification

In order to classify our collection of tweets, we applied deep learning by training a Convolutional Neural Network (CNN) as defined in (Kim 2014). Our choice is due to their results, and the pioneering in using such approach for classifying natural language. The results of the model presented in (Kim 2014) outperformed traditional methods, such as Support Vector Machine (SVM), in a variety of text classification tasks and since then it is widely referenced in the literature. Another motivating factor was the automatic learning capability that deep learning has by incorporating improved learning procedures that make use of computing power and training data, working well on large sets of data (Ain et al. 2017). In a nutshell, the CNN architecture comprises four layers. The first layer converts words into vectors of low-dimensional representation called *word embeddings*. The second layer applies a series of convolutions over these word embeddings to produce a feature map for each sentence. The third layer is responsible for filtering the most important features into one feature vector trough a max polling operation. The fourth layer applies the softmax function to classify sentences into labels. The Python code of the CNN implementation we used is publicly available [10] [11] [12], and it is designed to be executed on the top of TensorFlow[13], an open source software library for high performance numerical computation.

4. Experiments

We developed a few experiments with our CNN in order to find the most suitable classification model for our emotion categories. The CNN parameters we used were the same as in (Kim 2014) because their results were built using these parameters, and all the variations we tried did not provide significant difference on our results. Our experiments were focused on the input provided to the CNN, which is the training set. Given that our limited number of labelled tweets did not provide enough data for properly training the CNN, we tried different approaches for gathering enough training seeds for our emotion categories:

- Distant supervision (Go et al. 2009; Purver and Battersby 2012; Suttles and Ide 2013): we applied distant supervision and used the emotion-labelled electoral tweets provided by (Mohammad et al. 2015) as training seeds. This resulted in 2,575 seeds.
- Filtering by keywords: we analyzed samples of our dataset and defined keywords that were likely to represent emotions in a tweet. The process for obtaining our keywords set was the same as for labeling tweets for our gold standard. We randomly selected sets of tweets and identified specific keywords that indicated presence of emotions of our emotion categories. Afterwards we checked other samples for such keywords and verified that tweets containing them were likely to belong to the respective emotion category, we also confirmed the presence of such keywords in our gold standard. We then filtered the tweets by these keywords and considered them as training seeds. This resulted in 4,019 seeds. Keywords used for filtering can be seen in Table 3.

[10] https://github.com/cahya-wirawan/cnn-text-classification-tf
[11] https://github.com/cahya-wirawan/cnn-text-classification-tf
[12] https://github.com/dennybritz/cnn-text-classification-tf
[13] https://www.tensorflow.org/

Table 3. Keywords used for filtering training seeds for the CNN

Emotion	Keywords
anger	anger, fuck, fucked, pissed, lmaof, damm
disgust	disgust, disgusted, disgusting
fear	worried, worry, scary, scaring, scared, fear
sadness	sad, sadness, saddened
surprise	surprised, surprising, surprise, shocked, shocking

Table 4. Results for the generated CNN prediction models

Approach	Avg. Precision	Avg. Recall	Avg. F-measure
Distant Supervision	0,4049	0,2099	0,1087
Keywords	0,7348	0,718	0,6846
Hashtags	0,322	0,3014	0,2783
Dictionary-based	0,5666	0,2958	0,2722

- Filtering by hashtags (Mohammad 2012): we used emotion hashtags collected from (Mohammad 2012) to provide automatic labelling. Labelled tweets were used as seeds. This approach resulted in only 150 seeds.
- Dictionary-based filtering: we used a lexicon approach and filtered tweets with the emotion categories available in NRC (Mohammad and Turney 2013). Tweets in which one emotion prevailed were filtered and used as training seed. This resulted in 23,153 seeds.

For each approach, the CNN was trained and a prediction model was generated. In all of our experiments, training seeds for the ”none” category were chosen by selecting tweets that did not contain any of the following terms: a) defined keywords used as seeds (Table 3), b) emotional hashtags defined in (Mohammad 2012), c) emotion expressions labeled according to the NRC lexicon (Mohammad and Turney 2013). We randomly selected 1,000 tweets for the ”none” class. The test was always conducted against our labelled set of tweets. To improve our results, we did as in (Kim 2014) and loaded in our CNN pre-trained word embeddings for all the experiments. We chose the word embeddings corpus provided by GloVe [14] because it is extracted specifically from tweets. Following (Kim 2014), the use of pre-trained word embeddings is an approach commonly used to improve performance when the training set is not large enough. Incorporating the GloVe's embedding set in the CNN improved our results. General results of our models can be found in Table 4.

As we can see, distant supervision did not provide the best of the results. One explanation could be due to the peculiarities of our context, which includes words and expressions different than those of an electoral debate context. The approaches based on lexicon and hashtags did not provide good results as well. We noticed a very high level of absence of emotion hashtags in our dataset, which resulted in very few seeds, not enough to generate an accurate prediction model.

From all of our experiments, the one filtering by keywords provided the best results and therefore was the one used to generate our prediction model. We con-

[14]https://nlp.stanford.edu/projects/glove/

Table 5. F-measure for the model generated by filtering keywords

Emotion	anger	disgust	fear	sadness	surprise	none
F-measure	0,86033	0,6589	0,6280	0,7207	0,5932	0,6462

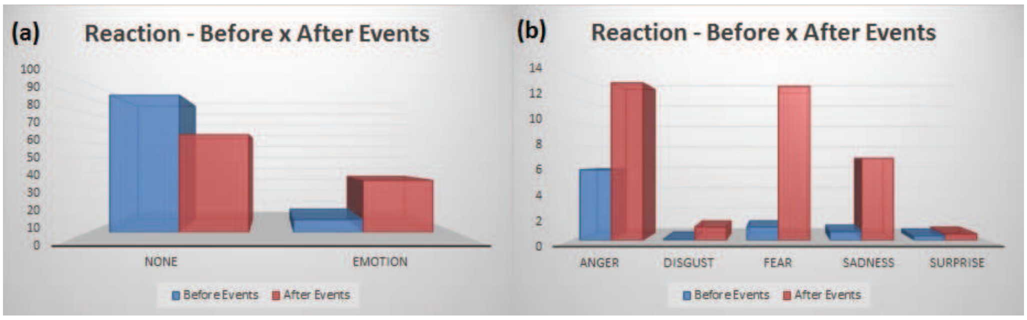

Figure 1. Tweets distribution before events and after events.

sidered our model reliable because it achieved average precision and recall above 70%, which we believe were good results taking into account results presented in (Suttles and Ide 2013; Purver and Battersby 2012; Mohammad et al. 2015). F-measure results for such a model can be seen in Table 5. It can be seen that the model's result for anger stands out along with the one for sadness. Remaining emotions have similar results, excluding surprise that performed below 60% but still close to the average.

5. Analysis

The first question we wanted to answer with our dataset was if there exists an emotion shift due to terrorism events. To answer this question, we compared emotion distribution before the events (BM and BL), and after them (AM and AL, respectively). Figure 1 depicts this comparison, where Y axis represents the percentage with regard to the total number of tweets of the respective dataset. All tweets are considered in Figure 1.(a), whereas only tweets with emotion are shown in Figure 1.(b). The first result observed was that before the events just about 8% of the tweets contained emotions from our emotion categories while after the events that number increased to around 25%. Furthermore, three emotions prevailed after the events: anger, fear, sadness. No significant changes were observed for disgust and surprise. Therefore, we conclude that there is indeed an emotional shift due to terrorism events.

The second question was whether different terrorism events raise the same emotional reaction. To answer this question we compared AM and AL in terms of emotion distribution. Figure 2 depicts emotion distribution for both events, where the Y axis represents, for each class (#prayformanchester and #londonbridge), the percentage of its total number of tweets distributed in emotion categories. Only tweets with emotion are shown. The results reveal that there are differences between these two events. While the event in London raised anger in the majority, the one in Manchester raised in the majority fear, followed by sadness.

A demographic analysis helped us understanding the differences between these two events. Figure 3 depicts emotion distribution by gender and age. Figure 4 shows gender and age distribution for both events. The Y axis represents, for each class (Gender

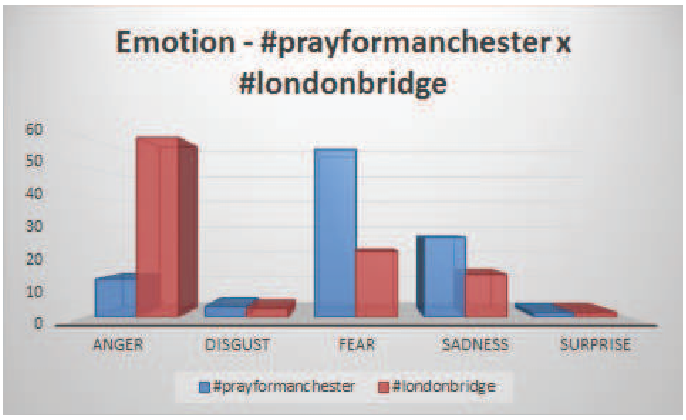

Figure 2. Emotion distribution for terrorism events.

Figure 3. Emotion distribution by Age and Gender.

and/or Age), the percentage of its total number of tweets per category for each dataset. In all graphs, tweets without emotion were not shown. The distribution of tweets by emotions for genders showed that Female users feel more fear and sadness compared to Male ones, who feel more anger instead. In addition, the distribution of tweets by emotion for ages shows that as the age increases, the feeling of anger increases proportionally. Fear, on the other hand, is higher for young ages, and it smoothly drops as age increases. No particular behavior was observed with regard to demographics for sadness. With these results, we distributed age and gender for both events and observed that there are indeed differences due to the concerned audience. In the Manchester event, the majority of tweeters are young women, i.e. the exact profile who mostly feels fear. This can be explained by the fact that Ariana Grande is very popular in this demographics. In the London event, such distribution showed that the majority of the tweeters were male middle-aged or older, i.e. the exact profile who feels anger. We believe that London

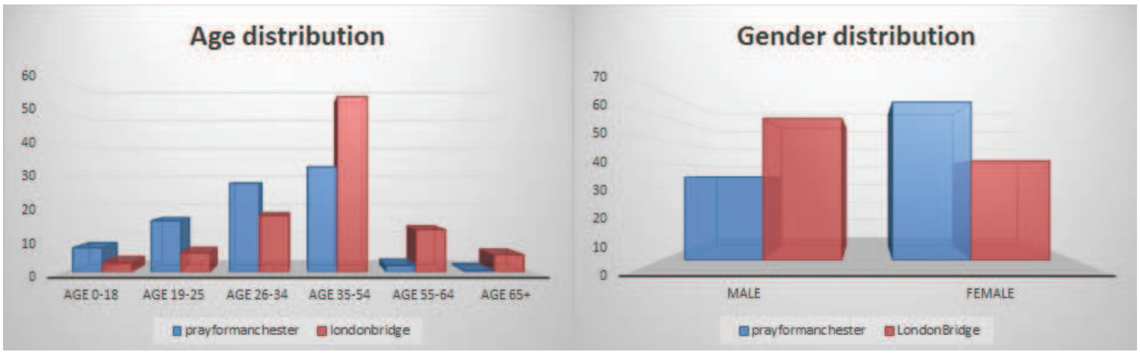

Figure 4. Tweet distribution by Gender and Age for both events.

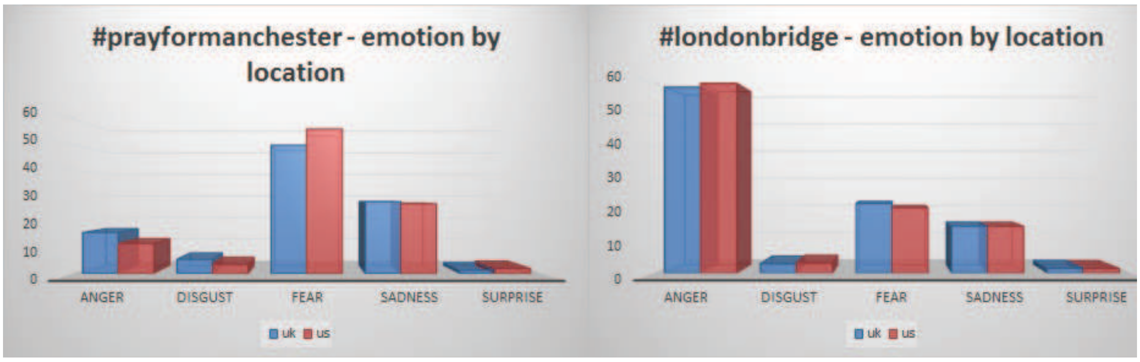

Figure 5. Tweet distribution by Location for both events.

bridge have affected the average London citizen who could be potentially at the location of the attack. Thus, we conclude that each terrorism event may raise distinct predominant emotions. Our hypothesis, to be confirmed, is that it is related to the people who see themselves as potential victims of a similar attack.

The third question was whether location has an impact on emotional reaction. To answer this question, we compared tweets from UK and US. The distributions for each country are depicted in Figure 5, where Y axis represents, for each class (UK and US), the percentage of its total number of tweets distributed in emotion categories. Only tweets with emotions are considered. For both locations, the distribution of tweets into emotion categories for both events did not show any noticeable variation. These findings indicate that location may not be an important factor as much as age and gender are.

6. Conclusion

Our work provided a study on the emotional reaction of twitter users to terrorism events. We addressed negative emotions and used deep learning approach for emotion prediction. Demographic data such as location, age and gender were extracted with help of available tools. Our results showed that when terrorism events occur, a shift of emotion towards anger, sadness and fear can be noticed. In addition, our demographic analysis showed that gender and age have influence on how tweeters react to terrorism events. Our data indicated that young Women tend to feel fear and sadness while Man in middle age and above tend to feel anger. Location did not provide any noticeable impact on the emotional reaction.

As contribution, we derive an emotion dataset in the context of terrorism and provided a CNN model that achieved good performance for emotions in our context. The questions we answered were a first step towards understanding the emotional reaction terrorism events raise on general population. We hope our work encourage further studies on social media focusing on terrorism, which we believe impact people in a complex emotional way. The data we provided might be used for further analysis and the results we reported might be used to better developing specific assistance programs for coping with terror. One opportunity is to improve our work by selecting similar terrorism attacks, as the differences of our targeted events might bring some noise to our analysis when comparing them. Another opportunity is to consider the location as indicated by georeferenced tweets and then study its possible impact. This because even if georeferenced tweets constitute a small set of the total, their information may be more accurate then the ones filtering by the location indicated in the users' profile.

References

[Ain et al. 2017] Ain, Q. T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B., and Rehman, A. (2017). Sentiment analysis using deep learning techniques: A review. *International Journal of Advanced Computer Science and Applications*, 8(6).

[Anderson 2005] Anderson, B. (2005). Imagined communities. *Chap*, 4(Hansen 1999):48–60.

[Bravo-Marquez et al. 2016] Bravo-Marquez, F., Frank, E., Mohammad, S. M., and Pfahringer, B. (2016). Determining word-emotion associations from tweets by multi-label classification. In *Proc. of the IEEE/WIC/ACM WI*, pages 536–539.

[Burnap et al. 2014] Burnap, P., Williams, M. L., Sloan, L., Rana, O., Housley, W., Edwards, A., Knight, V., Procter, R., and Voss, A. (2014). Tweeting the terror: modelling the social media reaction to the Woolwich terrorist attack. *Social Network Analysis and Mining*, 4(1):1–14.

[De Choudhury et al. 2016] De Choudhury, M., Jhaver, S., Sugar, B., and Weber, I. (2016). Social media participation in an activist movement for racial equality. In *Proc. of the ICWSM*, pages 92–101.

[Ekman and Friesen 1982] Ekman, P. and Friesen, W. (1982). Emotion in the human face system. *Cambridge University Press, San Francisco, CA,*.

[ElSherief et al. 2017] ElSherief, M., Belding, E. M., and Nguyen, D. (2017). # notokay: Understanding gender-based violence in social media. In *Proc. of the ICWSM*, pages 52–61.

[Fan et al. 2014] Fan, H., Cao, Z., Jiang, Y., Yin, Q., and Doudou, C. (2014). Learning deep face representation. *CoRR*.

[Gallegos et al. 2015] Gallegos, L., Lerman, K., Huang, A., and Garcia, D. (2015). Geography of emotion: Where in a city are people happier? In *Proc. of the WWW*, pages 569–574.

[Go et al. 2009] Go, A., Bhayani, R., and Huang, L. (2009). Twitter Sentiment Classification using Distant Supervision. *Processing*, 150(12):1–6.

[Hasan et al. 2014] Hasan, M., Rundensteiner, E., and Agu, E. (2014). EMOTEX: Detecting Emotions in Twitter Messages. *ASE BIGDATA/SOCIALCOM/CYBERSECURITY Conference*, pages 27–31.

[Horgan 2014] Horgan, J. (2014). *The Psychology of Terrorism, Second Edition*. Taylor & Francis Group.

[Kim et al. 2011] Kim, S., Bak, J., Jo, Y., and Oh, A. (2011). Do You Feel What I Feel ? Social Aspects of Emotions in Twitter Conversations. *NIPS Workshop*, pages 495–498.

[Kim 2014] Kim, Y. (2014). Convolutional neural networks for sentence classification. In *Proc. of EMNLP*, pages 1746–1751.

[Lerman et al. 2016] Lerman, K., Arora, M., Gallegos, L., Kumaraguru, P., and Garcia, D. (2016). Emotions, demographics and sociability in twitter interactions. In *Proc. of the ICWSM*, pages 201–210.

[Liu 2012] Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1):1–167.

[Lotan et al. 2011] Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., and danah boyd (2011). The arab spring— the revolutions were tweeted: Information flows during the 2011 tunisian and egyptian revolutions. *International Journal of Communication*, 5(0).

[Mitchell et al. 2013] Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., and Danforth, C. M. (2013). The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLOS ONE*, 8(5):1–15.

[Mohammad 2012] Mohammad, S. (2012). #Emotional Tweets. In *Proc. of the First Conference on Lexical and Computational Semantics*, pages 246–255.

[Mohammad and Turney 2013] Mohammad, S. M. and Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. 29(3):436–465.

[Mohammad et al. 2015] Mohammad, S. M., Zhu, X., Kiritchenko, S., and Martin, J. (2015). Sentiment, emotion, purpose, and style in electoral tweets. 51(4):480–499.

[Munezero et al. 2014] Munezero, M. D., Montero, C. S., Sutinen, E., and Pajunen, J. (2014). Are they different? affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE Transactions on Affective Computing*, 5(2):101–111.

[Purver and Battersby 2012] Purver, M. and Battersby, S. (2012). Experimenting with Distant Supervision for Emotion Classification. *Proc. of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 482–491.

[Sakaki et al. 2010] Sakaki, T., Okazaki, M., and Matsuo, Y. (2010). Earthquake shakes twitter users: Real-time event detection by social sensors. In *Proc. of the 19th International Conference on World Wide Web*, pages 851–860.

[Suttles and Ide 2013] Suttles, J. and Ide, N. (2013). Distant supervision for emotion classification with discrete binary values. In Gelbukh, A., editor, *Computational Linguistics and Intelligent Text Processing*, pages 121–136, Berlin, Heidelberg. Springer Berlin Heidelberg.

[Wan and Paris 2015] Wan, S. and Paris, C. (2015). Understanding Public Emotional Reactions on Twitter. *Proc. of ICWSM*, pages 715–716.

[Wang et al. 2012] Wang, W., Chen, L., Thirunarayan, K., and Sheth, A. P. (2012). Harnessing twitter ”big data” for automatic emotion identification. In *2012 International Conference on Privacy, Security, Risk and Trust*, pages 587–592.