Scientific Research Publishing

Open Journal of Social Sciences, 2018, 6, 142-158
http://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

#ISIS—A Comparative Analysis of Country-Specific Sentiment on Twitter

Sara Denise Ruhrberg, Giulia Kirstein, Tim Habermann, Janina Nikolic, Wolfgang G. Stock

Information Science, Heinrich-Heine-University, Düsseldorf, Germany
Email: Kirstein.ruhrberg@gmail.com

How to cite this paper: Ruhrberg, S.D., Kirstein, G., Habermann, T., Nikolic, J. and Stock, W.G. (2018) #ISIS—A Comparative Analysis of Country-Specific Sentiment on Twitter. *Open Journal of Social Sciences*, **6**, 142-158.
https://doi.org/10.4236/jss.2018.66014

Received: December 25, 2017
Accepted: June 16, 2018
Published: June 19, 2018

Abstract

The aim of this study is to analyze tweets on Twitter to the topic Islamic State regarding their positive and negative emotions by performing a sentiment analysis. People of different regions and cultures have a specific emotionality concerning the IS, not only in daily life but also in writing microblogs. With the help of sentiment analysis, the following question should be answered: "What are Twitter user's opinions on ISIS in different states worldwide?" For this purpose, a Python tool is developed that interacts with the Twitter Streaming API to retrieve Tweets that are IS-related, saving them with associated countries. Close to 500,000 tweets are collected by this tool over a period of nearly six weeks. Sentiment analysis of Tweets is made with a tool invented by Janina Nikolic. Results are normalized with additional, self-developed Python scripts and analyzed with Microsoft Excel and IBM SPSS. The results show that most of the Tweets in the countries have a negative attitude towards the Islamic State, and only a very limited set of states has a neutral or positive total sentiment in the results of this study. Sentiment is influenced by various factors, including political systems, geographic location and distance towards the area where the Islamic State is active and terroristic attacks.

Keywords

AFINN, ISIS, Terrorism, Sentiment Analysis, Data Analysis, Senti Strength, Emotions, Twitter, Hashtags, Emoticons

1. Introduction

People use various online social media networks to represent themselves by their social interactions. One possible network for this is Twitter. It is a microblogging service, enabling its users to publish short text-posts. The length of these posts is

limited to 140 characters. The social network was founded in 2006 using the name "twttr" and was renamed into today's "Twitter" only a short time later [1]. Posts on Twitter are called "tweets", along with the 140 signs. Users are allowed to upload a photo or a video to add to their tweet. According to Twitters own statements, the service has got about 317 Million active users worldwide [1] with each of them using it at least once a month. About 83% of active users are mobile users, only using Twitter on a mobile phone or a tablet. Twitter enables its users to follow other users, a way to subscribe to their posts. There are different ways to react to another user's post, such as marking it as a "Favorite", responding to the tweet or "Retweet" it. The last option is a way to share a tweet with your own followers. The amount of interactions of a tweet enables us to make a statement about its influence [2]. Hashtags are used to generate groups of tweets and to tag important keywords. New trends can be identified by monitoring frequently used Hashtags on Twitter [3]. Lots of data from Twitter can be used publicly. Not only the content of tweets is available, but a lot of additional data like language, hashtags and geographic coordinates.

Islamic State (IS)—also known as Islamic State of Iraq and Syria (ISIS) is a Sunni jihadist unrecognized state and militant group, which has been designated a terrorist organization by the United Nations (UN) and other individual countries [4]. Parts of Syria and Iraq are still in control of the ISIS militia; the front lines are continuously shifting. **Figure 1** shows the territories controlled by ISIS in March 2015. Sentiments of these affected states will be discussed in the results. The organization was founded in 1999 under the name of Jamaat al-Tawhid wal-Jihad. In 2004, it changed its name to Tanzim Qaidat al-Jihad fi Bilad al-Rafiday—also known as al-Qaeda in Iraq (AQI) [5]. From 2006 till 2013, the organization called Islamic State of Iraq (ISI), and then ISI renamed to Islamic State of Iraq and the Levant (ISIL, ISIS, IS) [6]. Since 2014, the organization is only

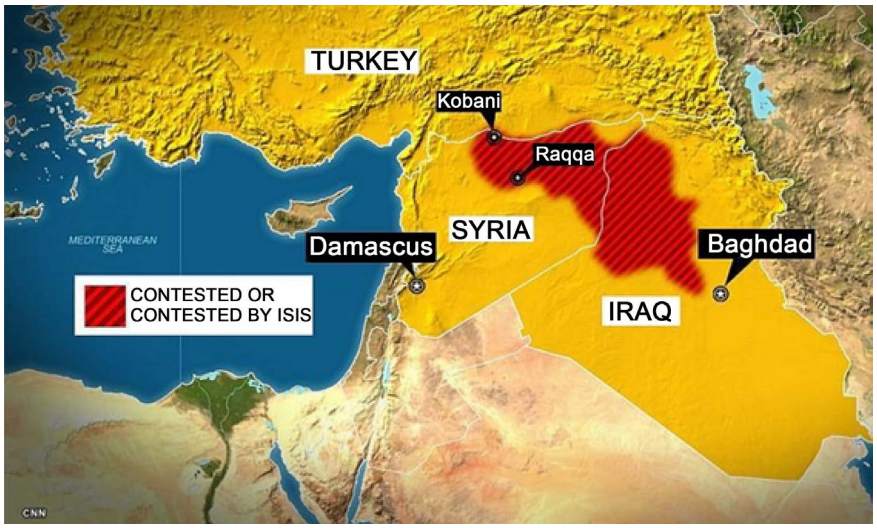

Figure 1. Map highlighting the countries of Iraq, Syria and Turkey called out are the cities of Mosul and Kobani. The area of ISIS controlled or contested territory is highlighted in red [9].

called Islamic State to show that the organization is acting across the borders.

The Islamic State is using Twitter for propaganda and terroristic support. It is the favorite social media tool of the terroristic organization especially for gaining support in western countries (those where Twitter is a favorite tool for marketing and online communication). This study aims to show the emotional echo on IS-related topics in those countries and other countries worldwide. Thus, a sentiment analysis using a self-programmed Python tool in addition to various statistic tools is performed. Sentiment analysis is based on diverse emotion lexicons like AFINN lexicon and a variety of methods, including SentiStrength.

2. Motivation

Sentiment analysis aims to determine the attitude of a person regarding a topic. Given a natural language text, e.g. tweets, it identifies whether the expressed opinion is positive, negative or neutral. In case of international comparisons of sentiment analysis, it is often not taken into consideration that there could be a fundamentally different distinct emotionality in countries though. For example, people of one country could react more emotional in general than people of another country, regardless of whether the topic is politics, sports or entertainment. This is not an unfounded assumption but underlies different studies about cross-country emotionality.

ISIS is known for its extensive and effective use of propaganda, especially by their social media strategies that redefine the use of propaganda in the 21st century [7]. Twitter is the favorite network of the Sunni militants in ISIS. The Islamic State is being supported by over 46,000 Twitter accounts worldwide are posting propaganda content [8]. Additionally, there are also Twitter accounts of news transmitters, accounts posting continuous negative attitudes, and of course accounts which are posting occasionally (**Figure 2**).

All these tweets together are forming sentiments on the Islamic State. The

Figure 2. Tweet of an ISIS supporting Twitter account for propaganda purposes.

present study addresses the following research questions: What are Twitter users' sentiments on ISIS and how are these sentiments distributed amongst countries worldwide? This will be answered by analyzing aggregated tweets regarding their sentiment.

In the present study, these opinions should be shown by analyzing tweets by country regarding the sentiment of these tweets calculated by various sentiment features (Words, Emoticons, Repeated Punctuation, Repeated Characters and Uppercase). The sentiments are clustered by positive or negative emotions. The evaluation is limited by certain hashtags mentioned later in methods (**Figure 3**).

State of Research

Magdy, Darwish and Weber [10] studied the antecedents of ISIS support in order to better understand the roots of the terror organization and its supporters. Therefore, Arabic tweets were collected and classified into pro-ISIS and anti-ISIS. Classification was done based on the "distinguishing language that signals current support or opposition for ISIS" [10]. Poblete *et al.* [11] investigated emotion-related differences across various countries on Twitter. They conducted sentiment analysis with English and Spanish emotion lexicons in order to define how happy the author of the tweet was. Besides, they examined the languages used per country and which hashtags were used by the users. A comparative study on explicit Twitter sentiment analysis determined "nine feature sets (41 attributes) which comprise punctuation, lexical, part of speech, emoticon, SentiWord lexicon, AFINN-lexicon, Opinion lexicon, SentiStrength method, and Emotion lexicon" [12]. Feature analysis was done by conducting supervised classification for each feature sets and continued with feature selection in subjectivity and polarity domain. By using four different datasets, the results revealed that AFINN lexicon and SentiStrength method are the best current approaches to perform Twitter Sentiment Analysis. In 2014, Berger investigated diverse hashtags which were

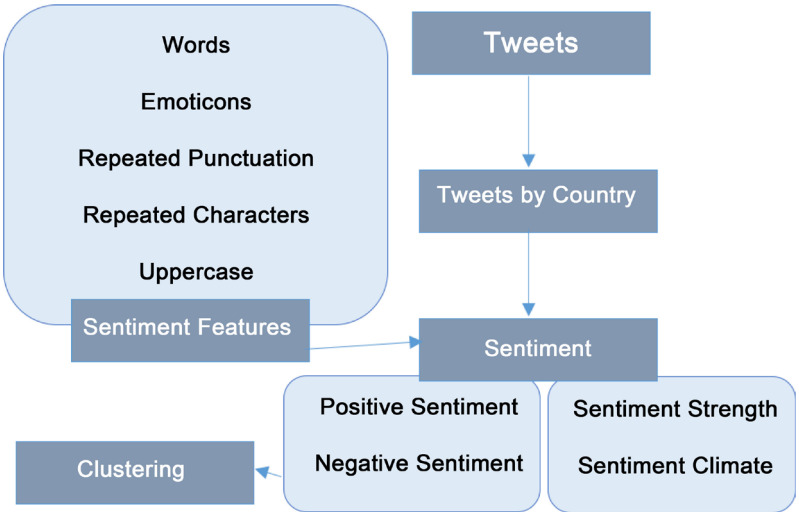

Figure 3. Research model shows the process of tweet processing with sentiment analysis tools.

used by ISIS to recruit new members, to promote targeted actions and wage war in the social networks. He discovered that ISIS also uses hashtags to draft a focus-group messaging and makes branding concepts [13].

3. Methods

3.1. Data Collection

Collecting a vast variety of different tweets concerning ISIS-topics was the first step. The social network Twitter is providing two different APIs for developers, a streaming API and a REST API. The streaming API allows to access a live stream of all tweets published, while the REST API makes old tweets accessible for developers. For unregistered developers, it only allows to collect tweets that are about seven days old, so for this project the REST API is not applicable. Further on, the official Twitter streaming API was accessed via Python applying a Python library called Tweepy to create a dataset for the sentiment analysis. The Twitter stream was filtered by hashtags, only saving tweets which contain at least one of the following tags: #isis, #is, #islamic_state, #Dawla, #Baqiyah. Selection of the last two hashtags was based on the ISIS-Twitter-Census from Berger and Morgan [8]. Berger and Morgan named additional hashtags which were not used in the present study, because they were written in Arabic letters and therefore could not be added to the Twitter search. There is a problem with using hashtags for a narrowly targeted data collection. Hashtags are not unique. Per definition, hashtags that are marked in a tweet should be important keywords that allow a categorization of the tweet. The topic of a tweet should be identifiable only by looking at the assigned hashtags. But in reality, users are not fully controllable and use hashtags at their own will. This leads to tweets with every word marked as a hashtag for example: "#Icecream #is #great". Looking at this example, the problem for the current data collection is already apparent. Since it is tagged with "#is", the filter will react and select the tweet as relevant for the dataset. But in fact, it has no relation to the Islamic State. This can influence the later sentiment analysis. Manually filtering all tweets tagged with "#is" is unfeasible. Deleting this hashtag from the set of monitored hashtags would be a bad decision as well, as the Islamic State identifies itself with these two letters. It was decided to keep the hashtag, but to conduct subsequent filtering in order to delete irrelevant tweets from the dataset. These filtering steps and their effect are explained later in this section. With every run of the program, two different types of datasets were saved in text files using comma separated values. Designed to be used for the sentiment analysis, the first dataset contains a unique ID, the follower count, coordinates (if included), time zone (if included), language and the actual text of every tweet. Most important for the upcoming sentiment analysis are the time zone and content of every tweet. Data-samples that were aggregated showed that only a minor percentage of ISIS-related tweets contained coordinates, so the time zone was chosen to assign a country to every tweet in the dataset. Simultaneously, a second dataset was saved, with every field that is in-

cluded in a tweet. This dataset can later be used as a backup or for further analysis to examine additional correlations. The program was run on a virtual Debian server for about five weeks, starting on 05-13-2016 and finishing on 06-20-2016. In this time, about 580,000 tweets containing at least one of the defined hashtags were saved by the tool. ReTweets, marked with "RT" by Twitter, were not considered since they do not represent the emotion or opinion of the user himself. This dataset had to be pre-processed before applying sentiment analysis.

3.2. Pre-Processing

A Python script was developed to simplify the pre-processing of the dataset. Not all of the 580,000 tweets included a time zone, so only tweets with a defined time zone were kept for further analysis. This reduced the dataset by about 50%, leaving 299,851 tweets. Using a list containing the time zones and their matching countries, the Python script replaced every time zone in the dataset by its associated country. Some time zones could not be replaced by a specific country, these tweets were deleted (e.g. GMT). As a result, the dataset now contained the unique ID, the language, the content of the tweet and the country. The follower count and coordinates were deleted in the process, as they were no longer considered necessary for the upcoming sentiment analysis. The first cleaning script left a dataset of 298,866 tweets from 113 different countries. This dataset still contained irrelevant tweets and needed additional filtering. Some tweets are not related to ISIS or any ISIS topics. The python script that was used for initial cleaning has been extended in order to detect relevant tweets and to sort out irrelevant items. As said earlier, especially the hashtag "#is" is problematic as it also reflects the conjugation of the English verb "to be". Some users tend to hashtag every word of their tweet, resulting in posts like "#icecream #is #great". All tweets containing "#is" or "#IS" are identified and double-checked by the Python tool. Only if a tweet contains additional terms that relate to ISIS, it is copied back into the final dataset. The following terms were identified to show a relation to ISIS-topics: "ISIS", "daesh", "IS", "islamic_state", "Dawla", "Baqiyah", "islamic", "jihad", "jihadist" "syria", "libya", "Libya" and "islam". The test conducted by the script was case-sensitive, so "IS" was added to the test as this term most probably always relates to ISIS. All other terms mentioned above are tested with various cases (e.g. "daesh", "Daesh", "DAESH") in order not to miss any relevant posts. After automated cleaning process, the dataset was scanned manually to ensure a clean dataset. To get significant results, every country with less than 100 tweets was deleted from the dataset, leaving 59 different countries with an overall sum of 246,454 tweets.

Table 1 shows the number of tweets after every step of cleaning up. With a bit more than 97,000 tweets, most tweets were published in the United States, followed by India and Serbia as shown in Table 2. On average 2298.204 tweets per country were found.

A lot of tweets were published in Malaysia, but manual screening of the dataset

Table 1. Number of tweets after several processing steps.

Step	Number of tweets in dataset
Raw data	580,000
Tweets with time zone	299,851
Tweets with country	298,866
Cleaned dataset	248,323
>100 per country	246,454

Table 2. Top-5-countries sorted by their number of tweets.

Country	Number of tweets in dataset
United States of America	97,118
India	38,208
Serbia	20,157
Great Britain	15,475
Japan	5458

showed that these tweets are mostly posted by bots, not recognized by the automated bot identification of the Python script. Thus, the results for Malaysia were treated separately and cannot be compared to other countries' results. After deleting the bot-tweets, Malaysia would have a total amount of only 66 tweets left (formerly 13,269 tweets including the bot tweets). All bot tweets were rated with zero by the sentiment analysis script, which would have led to a positive average sentiment for Malaysia. This will be further explained in the results-section. The dataset was now sufficiently cleaned. In an additional step, the actual content of every tweet needed cleaning too, before applying the sentiment analysis. Especially signs included in URLs (like "://" which could be recognized as an emoticon) and usernames can lead to false sentiments and to distorted results. In order to avoid this, a Python script to clean up the texts was developed. Usernames are replaced with "@USERNAME" and links are replaced with "URL". Also, the hashtag symbol has been removed from all tweet so that the words themselves can be identified. The language used for each tweet was downloaded from the Twitter API and saved alongside each post. So in a last step, all non-English tweet were translated into English to provide correct sentiment analysis. For translation, the Python libraries TextBlob and NLTK (Natural Language Tool Kit) were used in a small Python tool. It identifies the language of each tweet and provides an automatic translation that replaces the original text in the dataset.

3.3. Sentiment Analysis

All tweets have been analyzed on the granularity of a document level, so exactly one sentiment is assigned to a tweet. There are many different emotion lexicons and other possible features for conducting sentiment analysis—TextBlob (which was used for translation) for example also enables that, so it had to be decided

which one to use for this case. A comparative study on explicit Twitter sentiment analysis [12] investigated features which are good. They used 9 feature sets (41 attributes) that comprise punctuation, lexical, part of speech, emoticon, SentiWord lexicon, AFINN-lexicon, Opinion lexicon, SentiStrength method, and Emotion lexicon. Feature analysis was done by conducting supervised classification for each feature sets and continued with feature selection in subjectivity and polarity domain. By using four different datasets, the results reveal that AFINN lexicon and SentiStrength method are the best current approaches to perform Twitter Sentiment Analysis. AFINN is a list of English words rated for valence with an integer between -5 (negative) and +5 (positive), specifically constructed for microblogs. The words have been manually labelled by Finn Årup Nielsen in 2009-2011 [14]. SentiStrength is a lexicon-based classifier that uses additional (non-lexical) linguistic information and rules to detect sentiment strength in short informal English text. For each text, the SentiStrength output (for both version 1 and version 2) is two integers: 1 to 5 for positive sentiment strength and a separate score of 1 to 5 for negative sentiment strength. A neutral text would be coded as 1, 1. Two scales are used because even short texts can contain both positivity and negativity and the goal is to detect the sentiment expressed rather than its overall polarity [15]. While SentiStrength is more complex and already includes an emoticon lexicon, a negation lexicon etc., it is also a complete program and less easily to adjust. For this reason, AFINN has been chosen, but extended with further lexica and rules. So, all in all the following lexica have been used: the original AFINN emotion lexicon, an emoticon lexicon, a negation lexicon as well as a lexicon for booster words, like "very, totally" etc., and a lexicon for phrases, using SentiStrength as a model.

Condensed, the tool searches the text for specific sentiment keywords and emoticons, comparing them to a lexicon. Depending on which words, emoticons, repeated symbols and punctuation-marks are included in a tweet, the tool assigns negative and positive values. Looking at an example for words, four cases exist how sentiment is calculated. If a sentiment word has a negation word immediately in front, the value of the sentiment word is inverted. So, if "happy" has a positive value of 3, "not happy" gets a value of −3. A similar case is that you have a booster word between a negation and a sentiment word, so "not very happy" would also get a negative value. The third case is the combination of a sentiment word and a booster word in front, so "very happy" gets a boost value additionally and a total sentiment value of 4 instead of 3. The last case is simply a sentiment word with its own value. In addition to the word sentiment, emoticons and emojis are often used in tweets. While emoticons originated as text, and even today encompass both text (*i.e.* a simple form of ASCII art) and actual pictures, emoji have always been just pictures. Because of this, you could argue that emoticons are not just a subset of emoji, since they also have a text version, and indeed in Japan this text version has its own expanded symbols known as kaomoji (face letters). While it was important to include most of the emoticons

in the emoticon lexicon because of their facial expression, also a few emojis are counted that show no face, but also express a clearly emotion. For example, hearts in different colors or the party popper emoji. If simple emoticons appear one time, their value is added up to the whole emoticon sentiment for a tweet. If the same emoticon appears more often than one time, a booster of 0.5 is added for positive emoticons and −0.5 for negative emotions. The reason is that several people use these emoticons excessive, for example "haha, so funny!". But their emotionality is not really five times stronger as if they would just use one emoticon. When it comes to lengthened ASCII character emoticons like ":-))))", they also get a booster value of 0.5. The final tweet sentiment is calculated from text sentiment plus emoticon sentiment, where the text sentiment is normalized before.

All calculated values sum up to a final sentiment that is assigned to the tweet. The tool saves separate sentiments for text, emoticons and repeated symbols in addition to an overall sentiment for every tweet. To be suitable for further analysis, the computed sentiments need normalization.

3.4. Normalization

Specifically, for this dataset, an additional cleaning step was necessary. As some tests of the collected data showed, there are bots publishing posts with ISIS-hashtags. These sentiments should not be part of the results, since they do not represent human behavior and emotions. The sentiments for all bot-tweets that could be identified were set to zero, to ensure they are not represented in the results. Following this first step, two additional small scripts were run to edit the results. First, all sentiments were normalized using two factors which represent the strength of the sentiment for every country for positive and negative sentiments. These normalization factors for each country for positive and negative sentiments were elaborated by J. Nikolic in an earlier study. If people in a specific country are more emotional on Twitter than people from other countries, these factors will reduce this effect to generate comparable values. Sentiment values from countries with lower emotionality on Twitter are multiplied with a higher factor while sentiment values from countries with high emotionality are multiplied with a lower factor in this study. There are two different factors used, one for positive and one for negative sentiments. For example, the positive normalization factor for China is 1.123, while for United Arab Emirates it is 0.919. Normalization will therefore reduce German sentiments by about 8.1% while boosting Chinese values by over 12%. Equation (1) is the normalization formula to normalize all sentiments on the interval from −5 to +5 using the maximum absolute value of all tweets to get presentable graphs. This reduces the effect of rogue results and makes graphs and tables easier to read.

$$\left(\frac{5\times\left(sentiment+\max\left(abs\left(sentiment\right)\right)\right)}{2\times\max\left(abs\left(sentiment\right)\right)}-2.5\right)\times 2 \tag{1}$$

3.5. Post-Processing

Every tweet with its sentiment was imported into IBM SPSS 23 for further analysis. To compute an average value for every country in the dataset, results were grouped by country and sorted by number of tweet per country. The results of data collection and the sentiment analysis will be shown in the following chapters.

4. Results

The results show, that referring to the research questions "What are Twitter user's sentiments on ISIS?" and "How are these sentiments distributed amongst countries worldwide?", people from different countries are tweeting with different emotions about ISIS-related content. Overall, a negative attitude (sentiment −0.788323618 with a standard deviation of 0.412352909) in the tweeting world prevails regarding the Islamic State. All the evaluated countries have shown a negative attitude towards ISIS. There are two outliers from this statistic: the most negative tweeting country Serbia and Malaysia as the most positive tweeting country.

Figure 4 shows the distribution of Twitter sentiments that deal with the topic ISIS as a heat world map. Land sections that are marked white show no tweet on the topic in the period of tweet-acquisition. Clearly visible is that the color yellow-orange predominates. Due to the massive number of bot-tweets from Malaysia (see method-section) the natural most positive tweeting country is Japan with 5,458 evaluated tweets and a normalized total sentiment of −0.532935047. The most negative tweeting country is Serbia with a normalized sentiment of −2.3343716 and 20,157 evaluated tweets.

Unfortunately, only little and therefore not meaningful data could be collected from the direct affected environment of the Islamic State. But the existing meaningful data show at least a vague (negative) trend (Table 3 and Figure 5).

To determine the sentiment concerning ISIS, the most positive and negative tweeting countries further analysis steps were made. Table 4 and Figure 6 show the most positive tweeting countries with the total number of tweet evaluated and the final sentiment for each country. Malaysia is leading the list with an average sentiment of −0,005. All other countries included in Table 4 have a negative average sentiment, lower than −0.5. Figure 6 illustrates that the majority of positive tweeting countries is located in the eastern world.

Exemplary some of the most positive tweets of the most positive tweeting countries were selected to illustrate the attitudes. Tweets are shown as published originally, without the replacement of usernames and URLs which was done for sentiment analysis as described earlier.

Malaysia

- "Going to sleep with a smile on ma face #IS #AA [...]". (Unnormalized Sentiment: 1.015502355)
- "From ctrlsec0 Targeted #ISIS accounts\n [...]". (Unnormalized Sentiment: 0)

Italy

- "@Pontifex_it Paradise of bribes a strong word @rosariocrocetta the #ISIS

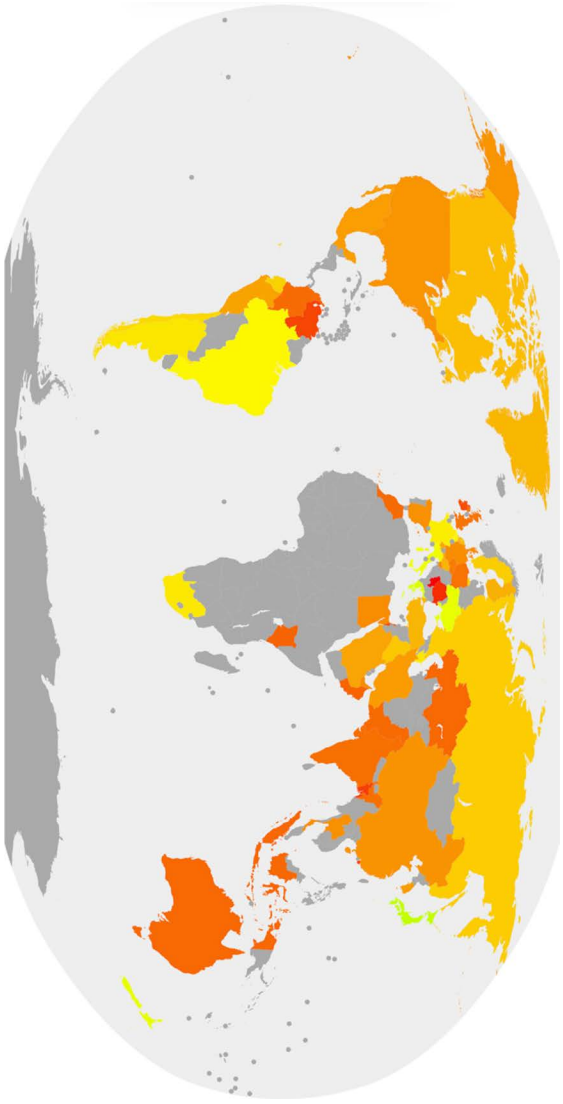

Figure 4. Heat map with ISIS-related sentiments in average for each country. The normalized scale covers the range from very negative (−5, illustrated red) to very positive (5, illustrated green). Countries marked grey have no sentiment in this study. Heat map was created using Gunnmap (http://gunnmap.herokuapp.com/).

Table 3. Countries directly affected by ISIS with their sentiments.

Country	Number of evaluated tweets	Sentiment
Iraq	3673	−0.71
Iran	416	−0.80
Oman	351	−0.96
Turkey	757	−0.75
United Arab Emirates	229	−0.84
Saudi Arabia	579	−0.78
Azerbaijan	166	−0.65
Egypt	817	−0.84

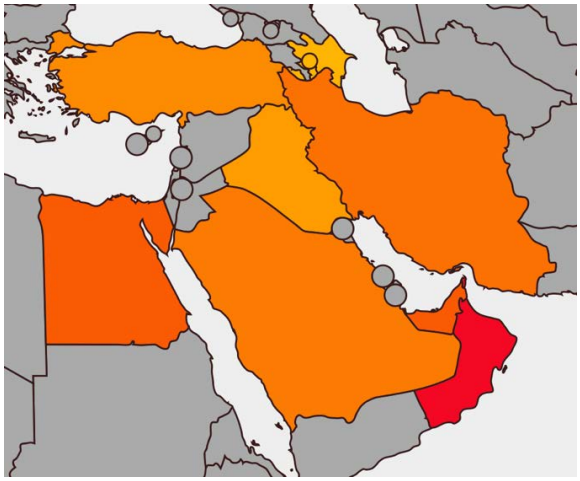

Figure 5. Heat Map with sentiments for each country located around ISIS controlled area.

Table 4. Top 10 most positive tweeting countries. Malaysia was ignored for further analysis due the amount of bot-tweets.

Country	Number of evaluated tweet	Sentiment
Malaysia*	13,269	−0.005
Japan	5458	−0.53
Greece	2510	−0.58
Ukraine	479	−0.59
New Zealand	104	−0.59
Italy	2504	−0.60

Arsonists as a strong accusation" (Unnormalized Sentiment: 1.208100151)

- "With @SassuoloUS in the Europa League Nek potr finally go international. @NekOfficial Great! #Neroverdi #ISIS #DioTifaTrotta" (Unnormalized Sentiment: 0.89678179)

Japan

- "Futile to try to win "hearts &amp minds" when they are twisted by centuries of religious dogma. Send them to paradise! #Fallujah #ISIS" (Unnormalized Sentiment: 1.670636901)
- "@janimine LOL neither the Quran nor Moh inspire #israelTerrorism and #ISIS are mercenaries recruited by israel-usa-saudi LOL" (Unnormalized Sentiment: 1.193312072)

Greece

- "@derStandardat Not only the West. Nearly all of Islam is really enthusiastic about the #Is... #Islamzombies #Grne #SPD #CDU #AfD #Spiegel" (Unnormalized Sentiment: 0.95892203)
- ". @DanieleRaineri: Telegram of the Islamic State #ISIS immediately created a channel to sing praises to the #Orlando martyr #Mateen" (Unnormalized Sentiment: 0.719191523)

Ukraine

- "@worldonalert Hahaha. sure. You are good at this. #Iraq #ISIS #FAIL #ProgandaScheisse" (Unnormalized Sentiment: 1.173886061)
- "#SAA has to be very careful. #IS moving fighters from all fronts to defend #Tabqa and #Raqqa. Time to secure supply line." (Unnormalized Sentiment: 0.978238384)

New Zealand

- "Possible that #ISIS will be finished by the end of #Ramadhan Time to celebrate." (Unnormalized Sentiment: 1.06456044)
- "The latest DnNet Daily! [...]". Thanks to @22SCN @KITTYGATTY @CharitySecurity #terrorism #isis" (Unnormalized Sentiment: 1.06456044)

As mentioned in the method-section most of the tweets from Malaysia were written by bots (see second exemplary tweet above). Those tweets were marked with the sentiment 0. When deleting all the tweets of the newly identified bots

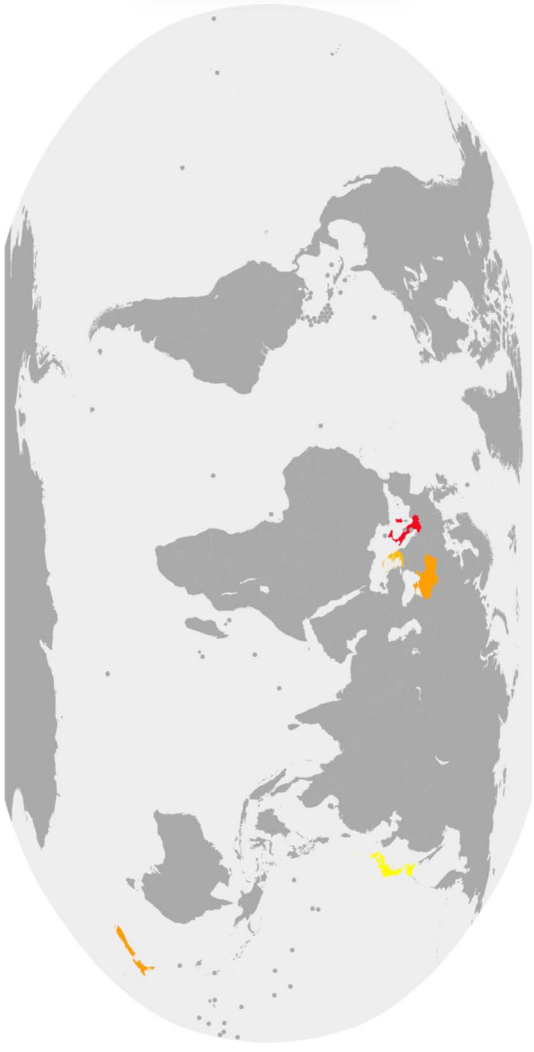

Figure 6. World map showing average sentiments for the top 5 most positive tweeting countries. Yellow represents more positive values while orange and red represent negative average sentiments.

the average sentiment would decrease from -0.0053 to -1.06. The multitude of neutral tweets makes ensure that the sentiment on country level becomes much more positive.

Table 5 and Figure 7 show the most negative tweeting countries. The country Serbia is leading the list with negative average sentiments. Figure 7 illustrates that the negative tweeting countries are located all over the world.

Exemplary some of the most negative tweets of the five most negative tweeting countries were selected to illustrate the attitudes.

Serbia

- "#China &amp #Russia Join forces to Destroy #ISIS while U.S sucks its thumb \n [...]". (Unnormalized Sentiment: −2.989590411)
- "#ISIS kills 250 women because they refuse to become their sexual slaves \n [...]". (Unnormalized Sentiment: −2.760872264)

Romania

- "VDO:127 People killed in #Iraq Bloody Tuesday: Iraqi Body Count Says \n#ISIS #Iraq #Baghdad\n [...]". (Unnormalized Sentiment: −2.956.339.152)
- "#ISIS Tortures #Children Like Men after Forcing Them to Confess to False Charges \nSee: [...]". (Unnormalized Sentiment: −2.878728414)

Bangladesh

- "Retweeted SITE Intel Group (@siteintelgroup):\n\n"Amaq Agency of #ISIS reported that ISIS fighters killed a... [...]". (Unnormalized Sentiment: −3.873626374)
- "IS-claimed Yemen bombings kill 41.\nRead more: [...] #Yemen [...]". (Unnormalized Sentiment: −2.95133438)

Venezuela

- "Retweeted Al-Masdar News (@TheArabSource):\n\nLebanese Army destroys another #ISIS sleeper cell... [...]". −3.348214286
- "RT: @rt_com:LEAKED: #ISIS covers fighters hefty bills in Turkish hospitals [...]". (Unnormalized Sentiment: −2.710459184)

Hong Kong

- "Bombs killed 100 in Syrian coastal cities of Jableh &amp Tartous. #IS claimed responsibility. [...]". (Unnormalized Sentiment: −3.019361591)
- "Against #IS. #SAA loss 5 Tanks in a day. After 2 Tank Destroyed in #Palmyra. #IS Destroy 2 more tanks in Qaryatayn. [...]". (Unnormalized Sentiment: −2.904683378)

Table 5. Top 10 most negative tweeting countries.

Country	Number of evaluated tweet	Sentiment
Serbia	20,157	−2.334371576
Romania	642	−1.219271532
Bangladesh	255	−1.108669635
Venezuela	582	−1.084986597
Hong Kong	176	−1.078643903

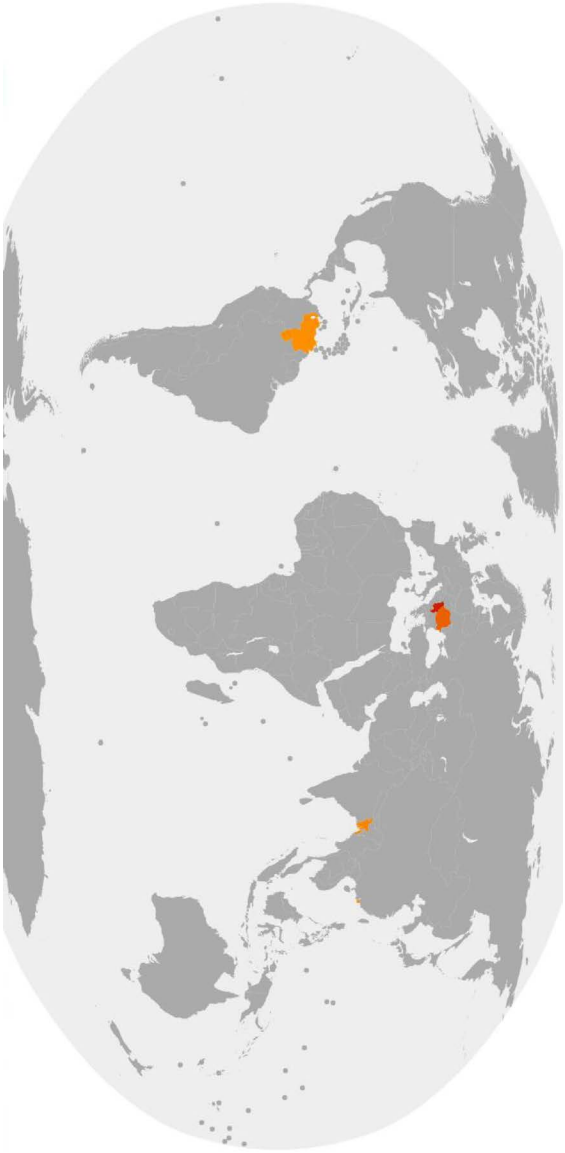

Figure 7. World map showing average sentiments for the top 10 most negative tweeting countries. Dark red represents more negative values, orange and yellow represent less negative sentiments.

5. Conclusions

The aim of this study was to investigate possible differences between countries according to the emotionality of tweets related to ISIS. Twitter and other social media networks are commonly used by terrorist organizations such as the Islamic State to promote their targets and to gain supporters. This paper shows how people in different countries think about the Islamic State and which emotions they have concerning actions of the terror militia. The world map in **Figure 6** shows that Russia and Southeast Asia are the most triumphant among the countries about the Islamic State. This could be due to the fact that Russia and Southeast Asia are far away geographically from the happening in the Middle

East and thus are not directly affected. Likewise, Russia and Southeast Asia are largely excluded from military action. In contrast to Vietnam, which is among the most positive countries, the sentimental values of Hong Kong show the most negative. In Hong Kong, there is no news control, as in other parts of Southeast Asia. Israel is also one of the countries with the most negative sentiment values. This could be because Israel is at the center of action and therefore is also affected by terrorism. The direct contact between military events and the people leads to the conclusion that people's tweets have more emotional values in contrast to an indirect contact by the news.

The results show that there is a country-specific emotionality based on tweets. To compare the sentiment of different countries according to a certain topic, this basic country-sentiment has to be subtracted out. "How would the development of a necessary normalization look like?" The normalization has to be made separate for positive and negative tweets because of a different positive and negative country-sentiment. The normalization factor is the sentiment mean about all countries divided by the sentiment mean of the particular country and has to be multiplied by the sentiment values of topic related tweets. There are countries that are more emotional in general (positive and negative), countries which are less emotional on principle and countries where only the positive or negative sentiment is strong.

6. Future Work

In this research, the data were only collected in a short period of time. For the future, it would be advisable to extend the period for having more data as a basis. In this study, it is noticeable that users use the word "IS" both for "Islamic State" or the verb "is". So, tweets should be manually searched for ambiguous hashtags to avoid these errors. It is also useful to expand on additional social media platforms, for example Facebook or/and Instagram. Other research areas could be that only certain countries are examined and compared. Regarding pre-incident events, they may have influenced the social media users.

References

[1] Sagolla, D. (2009) How Twitter Was Born. 140 Characters: A Style Guide for the Short Form.
http://www.140characters.com/2009/01/30/how-Twitter-was-born/
https://doi.org/10.1002/9781118258002

[2] Kietzmann, J.H., Hermkens, K., McCarthy, I.P. and Silvestre, B.S. (2011) Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media. *Business Horizons*, **54**, 241-251. https://doi.org/10.1016/j.bushor.2011.01.005

[3] Ma, Z., Sun, A. and Cong, G. (2012) Will This #Hashtag Be Popular Tomorrow? *Proceedings of the* 35*th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Portland, Oregon, 12-16 August 2012, 1173-1174.
https://www.researchgate.net/profile/Aixin_Sun/publication/254464722_Will_this_hashtag_be_popular_tomorrow/links/54449f120cf2a6a049ab157e.pdf

https://doi.org/10.1145/2348283.2348525

[4] Al-Ibrahim, F. (2014) Why ISIS Is a Threat to Saudi Arabia: Wahhabism's Deferred Promise? http://english.al-akhbar.com/node/21234

[5] Gambill, G. (2004) Abu Musab Al-Zarqawi: A Biographical Sketch. *Terrorism Monitor*, **2**, 1-4. https://jamestown.org/program/abu-musab-al-zarqawi-a-biographical-sketch-2/

[6] Dearden, L. (2014) Isis vs Islamic State vs Isil vs Daesh: What Do the Different Names Mean and Why Does It Matter. The Independent. http://www.independent.co.uk/news/world/middle-east/isis-vs-islamic-state-vs-isil-vs-daesh-what-do-the-different-names-mean-9750629.html

[7] Stone, J. (2014) ISIS Attacks Twitter Streams, Hacks Accounts to Make Jihadi Message Go Viral. International Business Times. http://www.ibtimes.com/isis-attacks-Twitter-streams-hacks-accounts-make-jihadi-message-go-viral-1603842

[8] Berger, J.M. and Morgan, J. (2015) The ISIS Twitter Census: Defining and Describing the Population of ISIS Supporters on Twitter. The Brookings Project on U.S. Relations with the Islamic World. https://www.brookings.edu/wp-content/uploads/2016/06/isis_Twitter_census_berger_morgan.pdf

[9] Nwaohuocha, O. and Tuysuz, G. (2015) Turkey Steps up Attacks vs. ISIS, Arrests Almost 600 Terror Suspects. CNN Philippines. http://cnnphilippines.com/world/2015/07/25/Turkey-intensifies-attacks-vs-ISIS-arrests-almost-600-terror-suspects.html

[10] Magdy, W., Darwish, K. and Weber, I. (2015) #FailedRevolutions: Using Twitter to Study the Antecedents of ISIS Support. https://arxiv.org/pdf/1503.02401v1.pdf

[11] Poblete, B., Garcia, R. Mendoza, M. and Jaimes, A. (2011) Do All Birds Tweet the Same? Characterizing Twitter around the World. *Proceedings of the* 20*th ACM International Conference on Information and Knowledge Management*, ACM, New York, NY, 1025-1030. https://dl.acm.org/citation.cfm?id=2063724

[12] Koto, F. and Adriani, M. (2015) A Comparative Study on Twitter Sentiment Analysis: Which Features Are Good? *International Conference on Applications of Natural Language to Information Systems*, Springer International Publishing, 453-457. https://www.researchgate.net/profile/Fajri_Koto/publication/277957183_A_Comparative_Study_on_Twitter_Sentiment_Analysis_Which_Features_are_Good/links/55777b0008aeacff20004c65.pdf https://doi.org/10.1007/978-3-319-19581-0_46

[13] Berger, J.M. (2014) How ISIS Games Twitter: The Militant Group that Conquered Northern Iraq Is Deploying a Sophisticated Social-Media Strategy. The Atlantic. http://www.theatlantic.com/international/archive/2014/06/isis-iraq-Twitter-social-media-strategy/372856

[14] Nielsen, F.Å. (2011) A New ANEW: Evaluation of a Word List for Sentiment Analysis in Microblogs. http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/6006/pdf/imm6006.pdf

[15] Thelwall, M., Buckley, K. and Paltoglou, G. (2012) Sentiment Strength Detection for the Social Web. *Journal of the American Society for Information Science and Technology*, **63**, 163-173. http://www.scit.wlv.ac.uk/~cm1993/papers/sentistrength2_preprint.doc https://doi.org/10.1002/asi.21662