Telematics and Informatics 35 (2018) 136–147

Contents lists available at ScienceDirect

Telematics and Informatics

journal homepage: www.elsevier.com/locate/tele

Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis

Nazan Öztürk[a], Serkan Ayvaz[b,*]

[a] Department of Computer Engineering, Bahcesehir University, Besiktas 34349, Istanbul, Turkey
[b] Department of Software Engineering, Bahcesehir University, Besiktas 34349, Istanbul, Turkey

ARTICLE INFO

Keywords:
Sentiment analysis
Text mining
Big data
Refugee
Twitter

ABSTRACT

The use of social media has become an integral part of daily routine in modern society. Social media portals offer powerful public platforms where people can freely share their opinions and feelings about various topics with large crowds. In the current study, we investigated the public opinions and sentiments towards the Syrian refugee crisis, which has affected millions of people and has become a widely discussed, polarizing topic in social media around the world. To analyze public sentiments about the topic on Twitter, we collected a total of 2381,297 relevant tweets in two languages including Turkish and English. Turkish sentiments were considered important as Turkey has welcomed the largest number of Syrian refugees and Turkish tweets carried information to reflect public perception of a refugee hosting country first handedly. We performed a comparative sentiment analysis of retrieved tweets. The results indicated that the sentiments in Turkish tweets were significantly different from the sentiments in English tweets. We found that Turkish tweets carried slightly more positive sentiments towards Syrians and refugees than neutral and negative sentiments, nevertheless the sentiments of tweets were almost evenly distributed among the three major categories. On the other hand, the largest number of English tweets by a significant margin contained neutral sentiments, which was followed by the negative sentiments. In comparison to the ratio of positive sentiments in Turkish tweets, 35% of all Turkish tweets, the proportion of English tweets contained remarkably less positive sentiments towards Syrians and refugees, only 12% of all English tweets.

1. Introduction

Social media has become an essential part of people's daily routine nowadays. Social media and internet can be used for various purposes including advertisement, spreading political opinions and financial trends, obtaining user comments about products, spreading spams, spreading news (Alarifi et al., 2016). Social media creates virtual bonds between users, in which people express opinions and develop relationships through posts, comments, messages and likes. Social media allows people to share their thoughts, feelings, opinions with other people instantly and easily.

While social networks are commonly used for purposes such as communication, information, advertising and social events (Chianese and Piccialli, 2016), they can also be used for political purposes as people tend to share political opinions on these platforms. Politicians often use public opinions for designing their campaigns in order to reach their target audience in a more effective and efficient way. Also, it is possible to analyze from social media posts to extract information regarding their supporters. For instance, former president of the USA, Barack Obama's campaign team used #Obama2012 and #AskObama hashtags on Twitter

* Corresponding author at: Yıldız Mh., Çırağan Caddesi, Osmanpaşa Mektebi Sokak 4-6, Besiktas 34349, Istanbul, Turkey.
E-mail address: serkan.ayvaz@eng.bau.edu.tr (S. Ayvaz).

http://dx.doi.org/10.1016/j.tele.2017.10.006
Received 2 August 2017; Received in revised form 11 October 2017; Accepted 14 October 2017
Available online 23 October 2017
0736-5853/

to reach voters and effectively carry out his presidential campaigns.

Users on social media create user profiles as a projection of their lives and often openly share personal information about them. Many companies have begun to analyze the user information on social media to reach their customers and conduct market research studies to collect customer opinions. On the other hand, customers also often interact with companies on social media, create content and express their thoughts and opinions regarding their products and services and (Bianchi and Andrews, 2015). Moreover, social media has attracted many research studies from various research communities from sociology, psychology, marketing, and communication to computer science (Pang and Lee, 2008).

Twitter is one of the most popular social media sites. Users post 500 million tweets daily in average on Twitter (Crannell et al., 2016) and 80% of those tweets are posted from a mobile device (Carley et al., 2016). Furthermore, the number of active Twitter users exceeds 22% of the internet users in the world (Kayser and Bierwisch, 2016). This means that more than 342 million people are actively using Twitter as of September 2016 (Statistic Brain, 2016). These statistics reflect Twitter's global outreach and potential impact. In addition to having a global coverage of issues, Twitter provides a media platform that enables sharing opinions easily using various content forms including text, images, links unlike many other social media platforms. Moreover, providing near real-time access to public posts through the API makes Twitter a suitable platform for large scale near real-time opinion mining. Hence, we primarily concentrated on Twitter as the main data source amongst other social media for the analyses in this study.

1.1. Background and related work

As an active research field that has emerged recently, sentiment analysis is a discipline that extracts people's feelings, opinions, thoughts and behaviors from user's text data using Natural Language Processing (NLP) methods (Heimann and Danneman, 2014). Moreover, sentiment analysis is also known as opinion mining, with emphasis on text classification problem. Extracting sentiment information from web-scale text data can be very challenging and expensive task due to large amount of data (Fernández-Gavilanes et al., 2016). Historically, the Web did not have much subjective data. With the explosion of social networks after early 2000 s, people began to share their feelings and subjective opinions through social media. Owing to increasing global coverage and high impact of social media, opinions of people on some issues that are shared through social media can be significantly influencing.

Unlike classical data mining methods, text mining and sentiment analysis deal with unstructured data (Oza and Naik, 2016). As a predominant sentiment analysis technique, lexicon approach is an unsupervised method, in which the text data are classified into a set of predefined sentiment classes. Sentiment scores of the text are calculated based on a sentiment lexicon, which is a dictionary consisting of words and their corresponding sentiment scores (Sun et al., 2017). From data modelling perspective, lexicon based methods can be viewed as models for determining the polarity of the text data. Lexicon based approaches may not provide high performance results since the terms in text data may carry a different polarity than specified in the lexicon (Muhammad et al., 2016). This problem can be avoided by generating context specific lexicons so that term polarity problem can be minimized.

Recently, Twitter has also become popular venue among scientists to conduct research studies from various perspectives. For instance, Lee et al. investigated the Twitter usage behaviors of journalists (Lee, Kim and Sang, 2017). In another study, the tsunami warnings in Padang Indonesia and reactions among Twitter users have been examined (Carley et al., 2016). A study by Crannel et al. investigated Twitter usage behaviors of cancer patients (Crannell et al., 2016). They found that cancer patients describe and explain their feelings about their diseases openly and candidly on Twitter.

Based on 36 million tweets collected from Twitter, Wang et al. proposed a real-time sentiment analysis system for classification of political tweets during 2012 US presidential elections. Their model achieved 59% accuracy in predicting the sentiments of political tweets (Wang et al., 2012). In another research study, Park et al. investigated the propagation of the issues discussed during TV debates in the 2012 South Korean presidential election to conversations on Twitter and communication patterns of Twitter users in political discussions (Park et al., 2016). A recent study analyzed the political campaigns in India during general elections and aimed to measure the effects of the internet on the first-time voters (Ahmed et al., 2016). In another related work, Bollen et al. proposed an approach to predict the movement directions in Dow Jones Industrial Average closing values (Bollen et al., 2011). Their method provided 86.7% accuracy in predicting the daily up and down directions of Dow Jones closing values.

On the other hand, Cheong et al., offered a sentiment analysis model that provides useful graphical visualizations about potential terrorism scenarios based on public sentiment data collected on Twitter (Cheong and Lee, 2011). In a case study, Yates and Paquette have explored the use social media technologies and their impacts on knowledge sharing and effective decision-making during 2010 Haitian Earthquake (Yates and Paquette, 2011). The case study demonstrated an example that social media and knowledge management technologies can play a critical role in effective response to disasters by sharing emergency information. In another recent study, Kim and Hastak investigated the characteristics of social networks after a disaster by investigating the social media responses after the 2016 flood in Louisiana (Kim and Hastak, 2018). They observed that the disaster related information diffused differently on social networks during emergency situations. Similarly, Neppalli et al. investigated the use of sentiment analysis on Twitter for extracting information in emergency cases. They analyzed the Twitter users' sentiments in affected locations during Hurricane Sandy. They also presented location-based sentiment changes (Neppalli et al., 2017).

From public health point of view, the spread of infectious diseases has also been analyzed using social media data by several studies (Lampos et al., 2010; Corley et al., 2010; Culotta, 2010; Broniatowski et al., 2013; Elkin et al., 2017). Another study performing sentiment analysis on a health-related issue, investigated the opinions of Twitter users about water-pipe smoking using Twitter data and they found out that 59% of the tweets have positive sentiment about water-pipe smoking (Grant and O'Mahoney, 2016).

Wu and Shen proposed a sentiment analysis model for prediction of news popularity on Twitter (Wu and Shen, 2015). They investigated the characteristics of news propagation on Twitter and found that there is a correlation between news popularity and the retweeters' frequency of interaction with the news source. A recent study focusing on Japanese Twitter users investigated the information overload and user reactions. The authors found out that users tend to not to look at all the tweet feeds on their timeline but they still continue to follow new users (Sasaki et al., 2016).

Another interesting research study is published recently that analyzes the sentiments in video game chat messaging to better understand the users (Thompson et al., 2017). The authors proposed a lexicon-based approach in which a made-up language generated by users is taken into account in the lexicon. For instance "gg" stands for "good game" in the lexicon. They suggest that lexicon-based approaches are portable and the dictionaries can be updated easily when needed to apply the model on other datasets. Dictionaries are generated according to the text data that is being analyzed. Thus, those approaches give an opportunity to cover the dataset with a well-prepared dictionary.

Sentiment analysis in microblogging platforms has also been drawing attention from researchers focusing on different languages. Tellez et al. performed a sentiment analysis study in Spanish language and analyzed text transformations using Support Vector machines as classifiers (Tellez et al., 2017). From text classification perspective, Liao et al. performed sentiment analysis on 30,000 Chinese microblogs and they classified messages into negative, neutral and positive categories (Liao et al., 2016). Using movie data, Khan et al. examined the sentiments of 50,000 movie reviews (Khan et al., 2016).

1.2. Syrian civil war and its effects on social media

After Syrian civil war in 2011, millions of Syrians had to leave their country. They took refuge in many countries including not only countries that have border with Syria such as Turkey, Lebanon and Jordan but also the western countries such as USA, Canada and Germany. According to www.syrianrefugees.eu, European Union (EU) is another major refugee hosting region, which also provides humanitarian aid for refugees as of 2016.

According to the official records of Ministry of Interior Directorate General of Migration Management of Turkey, there are 2,582,600 refugees living in Turkey in 2016 (Çetinkaya et al., 2016). According to an article in US News[1], there are 450,000 Syrians refugees in Germany, which is the biggest number in Europe as of 2016. Situation of the refugees are deliberated all around the world especially in those host countries. In the article published in Al Jazeera[2], 4.85 million of Syrians moved abroad due to the civil war and also 6.3 million of people relocated within their country, Syria as of 2017. These resettlements have brought along new questions such as will the refugees able to adapt the host country, will they go back their country after war ends, will they learn host country's language, what kind of rights they will demand if they stay longer, etc. These questions are also being asked in social media. The refugee crisis and issues involving Syrian refugees have become hot topics around the world. Hashtags like #syrianrefugees, #refugeeswelcome, #syria, #syrian, #refugeecrisis are used in social media for these topics. People from many countries have stated their opinions using those hashtags.

There have been several research studies investigating issues involving Syrian refugees from various perspectives. In a research study (Vaz et al., 2017), integration and education problems of Syrian refugees in Canada are discussed. Vaz et al. investigated potential refugee habitats on suburban regions around big cities using Geographic Information Systems. Furthermore, another study investigated Post-traumatic stress disorders in Syrian refugee children in a German refugee camp and examined 96 Syrian children between age 0 and 14 (Soykoek et al., 2017). Moreover, the living conditions of Syrian refugees in Turkey are investigated in another study (Villasana, 2016). They reported that the refugees are working as immigrant farmers and it is difficult for them to access health care services.

As a different perspective from aforementioned studies, our goal in this research study is to investigate the public opinions and sentiments towards Syrian refugees using text data from Twitter. To best of our knowledge, this is the first sentiment analysis study that explores the sentiments towards Syrian civil war and refugee crisis in a comparative analysis using data from social media. Social media is considered as a good venue because a mass number of people from various backgrounds use social media to discuss the topic and people are less reluctant to share their sincere thoughts on social media due to anonymity.

2. Material and methods

In this study, Turkish and English tweets were extracted from Twitter and analyzed in detail. Since Turkey currently hosts the biggest number of Syrian refugees, the topics involving the Syrian refugees are frequently discussed on Twitter. Turkish tweets carry valuable information regarding public perception of a hosting country first handedly. Furthermore, English is the most commonly used language in the entire world. People commonly use English language for posting tweets about Syrian refugees. Therefore, we decided to analyze Turkish and English tweets about the refugees.

2.1. Data collection

As a part of data gathering process, all potentially relevant tweets were searched and extracted from Twitter programmatically by

[1] https://www.usnews.com/news/best-countries/articles/2016-12-19/countries-hosting-the-highest-proportion-of-syrian-refugees.
[2] http://www.aljazeera.com/news/2017/03/number-syrian-refugees-passes-million-170330132040023.html.

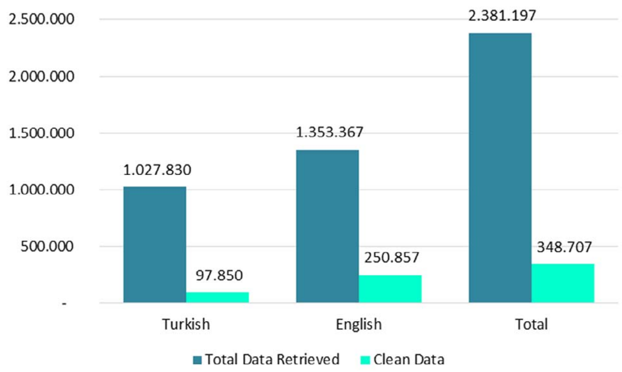

Fig. 1. The breakdown of tweets retrieved and cleaned.

using twitteR[3] package, written in R programming language. This comprehensive tweet search was conducted been 29/03/2017 to 30/04/2017. Consequently, the collection of related tweets was retrieved and saved in csv files. After all data were gathered, the files were combined and two different raw data files, one file per language, in preparation for the analysis process. The data in those files contained the tweet information along with user information posting the tweet. Posting dates were also substantial for the analyses. However, not all the tweets had the posting date information. The analyses in this study performed using the tweets with no missing values.

Turkish and English tweets were analyzed separately in the study. Thus, Turkish and English tweets were retrieved in two parts. For English Tweets 1353,367 tweets were gathered with the keywords "Syrian" and "refugee". For Turkish tweets, we used the corresponding Turkish keywords: "Suriyeli", "mülteci" and "multeci" for the search. In total, 1027,930 Turkish tweets were gathered.

2.2. Data preprocessing

There were a large number of duplicate tweets both in the Turkish and English data and tweets with missing values. Upon further investigation, we noticed that the iterative search using the Rest API of Twitter seems to cause these duplicates. Because Twitter API is not live stream and it provides a pool of tweets at a time window. Thus, the API is likely return a set of overlapping tweets if the requests were sent in short time intervals. During data cleaning, the retweets were also omitted as the aim of this study is to specify the sentiments or opinions of individuals and the retweets were not considered to reflect a new personal opinion. Therefore, we removed retweets from our analyses. As a result, 97,850 tweets, approximately 10% of retrieved Turkish tweet data and 250,857 tweets, nearly 19% of retrieved English data and 348,707 tweets in total, which make up 15% of all retrieved of data were kept for analysis after data cleaning process as shown in the Fig. 1 below.

In the preprocess part, all the uppercase letters were converted into lowercase characters. Punctuation, links and numbers were removed from the datasets and only words were kept for the analyses. After the preprocessing part, all tweets were split into tokens and those words were looked up in each sentence and scores were calculated. Then, the sentiment scores of the tweets were aggregated based on the token sentiment scores.

2.3. Development of Turkish sentiment analysis lexicon

For assessing the sentiment values of Turkish tweets, we had to develop a sentiment lexicon as there was no comprehensive lexicon available in Turkish. In the case of English tweets, we used RSentiment[4] package, which was developed in R programming language (Bose et al., 2017). RSentiment package contains an extensive sentiment dictionary, which helps analyze sentiment at the sentence level in English and provides a sentiment score. Thus, there was no need for developing a new dictionary in English.

To analyze Turkish tweets, we developed a broad sentiment lexicon, consisting of 5405 words. These terms were thoroughly chosen from the terms commonly used daily in Turkish language and were also among the most commonly used terms in the tweets from the retrieved data. These words were scored from −5 to +5 according to their sentiment value: −5 meaning very negative and +5 meaning very positive. In order to obtain an objective score, the terms in the lexicon were manually scored by 3 different graduate students, who are native Turkish speakers. Each term in the lexicon was scored separately by the individuals without knowing the score of the terms given by the other raters. To validate the reliability and accuracy of the final sentiment scores, we then manually reviewed the sentiment scores of the terms in the lexicon given by the raters. We observed that the sentiment scores of the raters were consistent for the majority of the terms in the lexicon. To ensure interrater reliability of the manually scored terms and to get a consensus on the terms that were scored differently by the raters, the averages of the group scores were calculated and assigned accordingly for each term in the lexicon.

[3] https://cran.r-project.org/package=twitteR.
[4] https://cran.r-project.org/package=RSentiment.

Since Turkish is an agglutinating language, terms may contain in various suffixes. These suffixes were not always detected by the algorithm used in the analyses. Thus, they were put into the dictionary as different words and scored accordingly. In order to avoid scoring issues due to typographical errors, some commonly misspelled words were also added to the lexicon. Furthermore, the swearing words used in the tweets were appended the dictionary as people on social media tend to use swear words frequently. City names, politician names and some abbreviations were included in the lexicon in order to preserve content integrity. The Turkish language lexicon that was developed in this study made available for download to interested researchers for use in prospective research studies.[5]

2.4. Calculation of term frequencies

Word clouds are useful tools for visually summarizing a large amount of text data. To visualize a mass amount of text data gathered in this study, word clouds were generated for both Turkish and English tweets. These word clouds were created by using wordcloud[6] package in R programming language, based on the frequencies of the words used in all retrieved tweets to determine sizes of words. The words in the cloud are located in accordance with their frequency.

3. Results

To compare sentiment scores of Turkish tweets and English tweets, we analyzed the sentiment scores of the tweets for each language separately. The sentiment scores of tweets were grouped into five categories: Very Negative, Negative, Neutral, Positive, and Very Positive.

3.1. Turkish tweet analysis results

We first analyzed the sentiment scores of Turkish tweets. Fig. 2 below shows the breakdown of Turkish tweets into the sentiment categories visually. Out of 97,850 Turkish tweets used in the analyses, 30,374 tweets were categorized as neutral, which constituted the largest category. With 8286 tweets, very positive category contained the least amount of tweets. However, the number of tweets in positive category was significantly higher than tweets in negative category, 25,970 versus 19,587 tweets, respectively.

Fig. 3 demonstrates the proportions of the sentiment categories below. Approximately 27% of all tweets were marked as positive and 8% of all tweets were very positive, 20% of all tweets were negative, 14% of all tweets were very negative and 31% of all tweets were neutral. When combining positives with very positive category and negatives with very negative category, 34,256 tweets, 35% of all tweets were considered as positive and 33,220 tweets, 34% of all tweets were considered negative, in total. The overall sentiment score of the Turkish tweets was slightly more positive than negative. However, the sentiment scores of the three main categories were very close to each other. With 35% of them being positive, 34% negative and 31% neutral, no sentiment category was dominating the others, proportionally.

Furthermore, we analyzed the sentiments of Turkish tweets over a period of time. Only the tweets that had the posting date information were used since not all retrieved tweets possessed the date information. In the Fig. 4, the sentiment categories of tweets are graphically shown based on the posting dates. In this figure, positive and neutral tweets are at the highest peak on 04/04/2017. On that day, a chemical gas attack allegedly made by Syrian government in city of Idlib was reported by news agencies around the world. On 07/04/2017, the US president Donald Trump ordered an air strike against Syrian government after the chemical gas attack. Negative tweets reached the peak point on 07/04/2017. On the same day, the number of negative tweets was nearly as much as the number of positive tweets, even though negative tweets were almost always significantly less by positive tweets for the rest of the time span.

On 07/04/2017, all sentiment categories reached their peak points. Second peak point for tweets regarding refugees was on 16/04/2017, which coincides with the date of the constitutional referendum for new presidential system, which was held in Turkey. The hike in the number of tweets stating opinions about Syrians and refugees could be associated with the referendum discussions.

3.2. English tweets analysis and results

For the sentiment analyses of English tweets, we used RSentiment, a package in R, which is designed for sentiment analysis in English language. It contains an English sentiment lexicon to score the words while calculating sentiments of the text data. RSentiment package also classifies texts into five sentiment categories, namely: very negative, negative, neutral, positive and very positive. The English tweet dataset contained 250,857 tweets. For each tweet an overall sentiment score was generated. Then, the tweets were grouped into the sentiment categories. Out of 250,857 tweets, 63,483 of them were classified as negative, 119,586 as neutral, 25,375 as positive, 36,903 as very negative and only 5509 tweets were classified as very positive.

The Fig. 5 visualizes the breakdown of sentiment analysis results of English tweets. We observed that neutral tweets dominated the sentiment categories. After neutral, negative tweets were the second in order and followed by the very negative tweets.

As shown in the Fig. 6, the proportions of the results composed neutral with the highest rate of 48% among all tweets, 25% as

[5] https://github.com/nazaan/Sentiment-Analysis.
[6] https://cran.r-project.org/package=wordcloud.

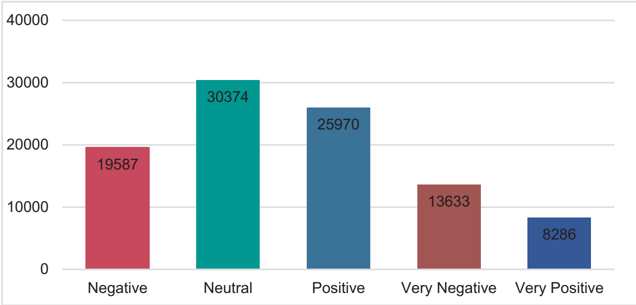

Fig. 2. Turkish tweets analysis results graph.

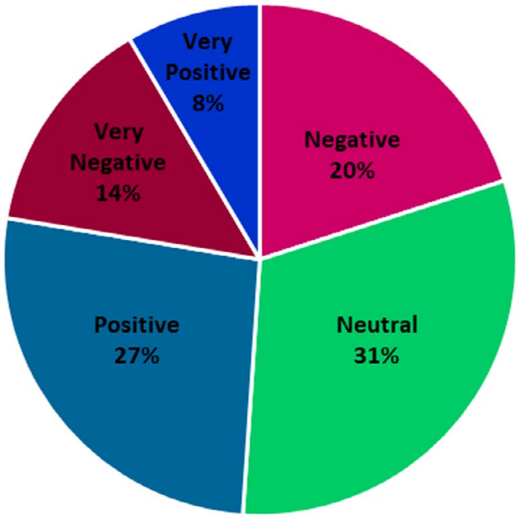

Fig. 3. Turkish tweets analysis results pie chart.

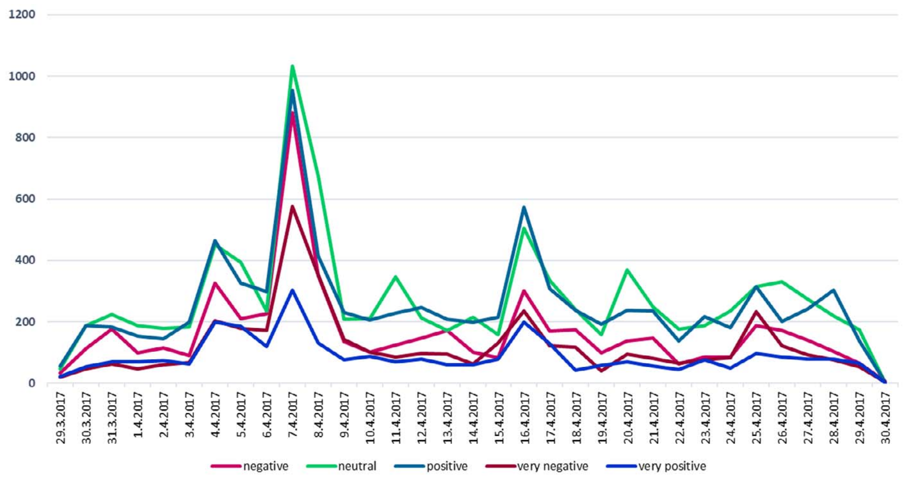

Fig. 4. Turkish tweets distribution graph by date.

negative, and followed by 15% as very negative, respectively. On the other hand, 10% of the tweets were categorized as positive and only 2% of all tweets were marked as very positive. Neutral tweets were distinctively dominating the results with 48% ratio but at the same time, a total of 40% of all tweets by combining negative and very negative categories were carrying negative opinions towards Syrians and refugees. Only 12% of the tweets were carrying positive sentiments. These results clearly show that English speaking community in Twitter posts tweets involving mostly neutral or negative opinions regarding the topics of Syrians and refugees.

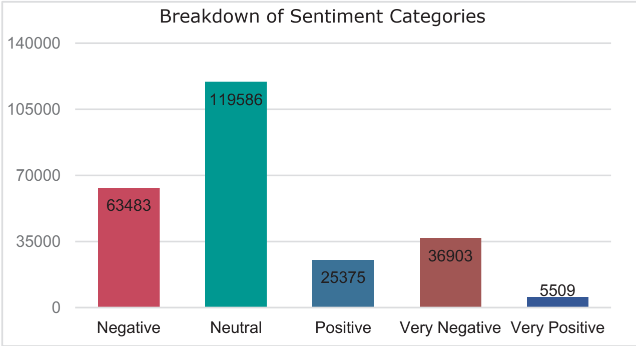

Fig. 5. English tweets analysis results graph.

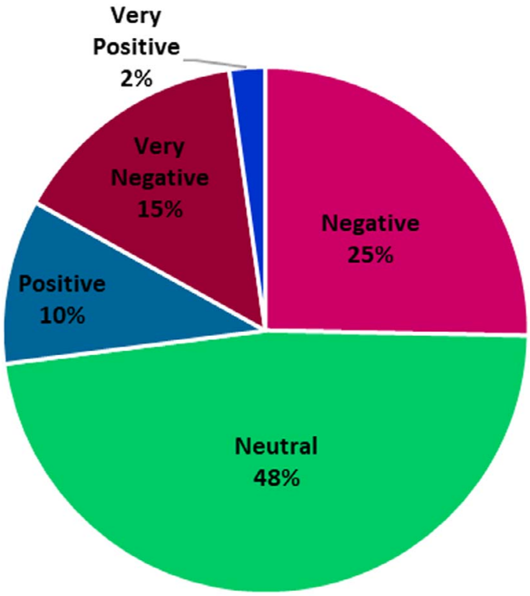

Fig. 6. English tweets analysis results pie chart.

Comparatively, the proportional sum of positive and very positive sentiment categories in English tweets was very low, 12% in English as opposed to 35% in Turkish tweets.

Fig. 7 above shows the sentiment score distribution of tweets by date for the tweets with the date information. In this figure, the neutral tweet category is always shown at the top position in comparison with other sentiments. Another observation from the figure is that very positive tweet category is always located at the bottom of the graph by a significant margin from the consequent category.

The same figure demonstrates that the number of tweets appears to be steady on the date of chemical attack, April 4, 2017. However, the tweets rapidly increased after April 7. On 15/04/2017, a vehicle carrying explosives hit a convoy near Aleppo and it caused 126 people's death. After this date, there was jump in the number of tweets. There was another big incident on April 27 th: a huge explosion happened near Damascus Airport. We think that the hike in the number of tweets on that day might be attributed to this incident.

3.3. Word clouds of frequent terms

We analyzed the word frequencies for both Turkish and English tweets about refugees using word clouds. Fig. 8 below demonstrates the word cloud results for Turkish tweets. The word "suriyeli", means "syrian" in Turkish, was the most common word among the terms in Turkish tweet data. Thus, it is shown as the boldest word in the cloud. Not surprisingly, it is followed by the word "suriye", meaning "syria" in Turkish.

As shown in the Fig. 9, the most common word was "syrian" in the word cloud that was generated from English tweets. Different from Turkish word cloud results, the second most frequent word was the word "refugee".

In Turkish tweets, the Turkish equivalents of the words "Turkish", "Turkey", "Iraq", "Syria", "Syrian", "Iran", "Arab", "Kurdish", "Afghan", "Russian", "Israil", "Russia", "Libya", "Kerkuk", "Idlib" were counted as frequent terms. As observed on the word cloud

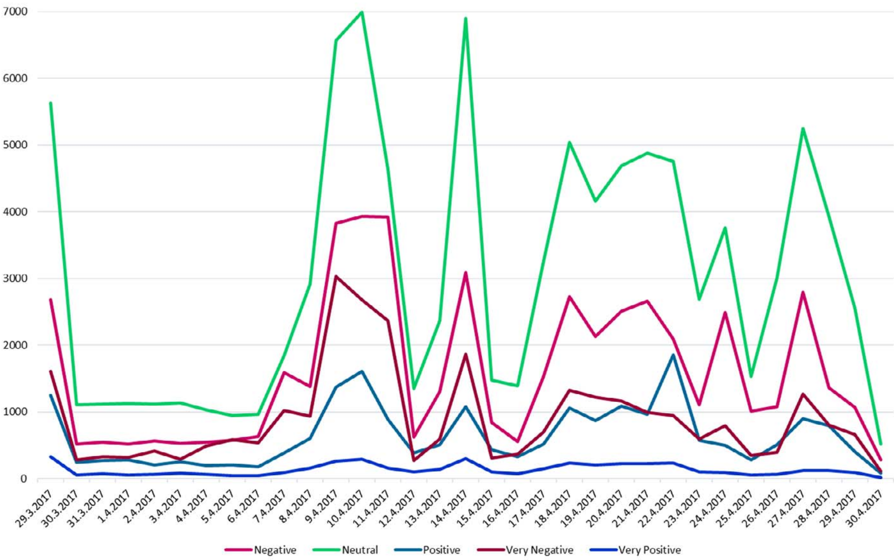

Fig. 7. English tweets distribution graph by date.

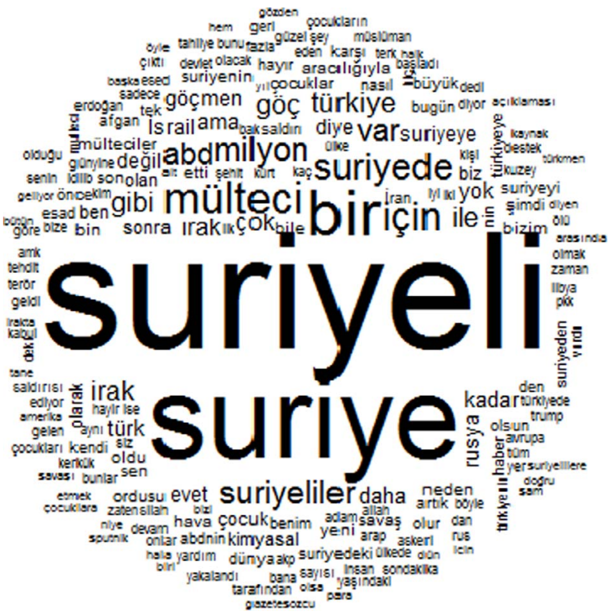

Fig. 8. Word cloud for Turkish tweets.

graphic of the frequent terms, the Turkish tweets often involved discussions regarding the geography, international relations and ethnicity issues. Considering geographical proximity of Turkey to the war taking place in Syria, the tweets in Turkish focused more on the news about the war, itself. Based on the content of the tweets, we observed that the people posting Turkish tweets appeared to follow the events in the war more closely. In English tweets, the words "Obama", "Trump", "Putin", "regime", "illegal", "president", "conflict", "crisis", "policy", "immigrant", "government" were counted as frequent. In contrast to Turkish Tweets, the English tweets seemed to focus more on the politics and legal side of the refugee crisis and slightly less on the details about the war in Syria.

Among the frequent terms in English tweets, there were no swearing words. In contrast, one swearing word was counted as a frequent term in Turkish tweets. There were some words frequently used in both datasets. These words were mostly related to the war. For instance, "bomb", "killed", "dead", "terror", "attack", "terrorist" and "hit". Apart from the war related words, "child" and "children" were counted as frequent. It appears that children of this war are taken into consideration and frequently thought for both English and Turkish speaking community. Regardless of the language of the community, children hold an important spot in the posts on Twitter regarding the war and refugee crisis.

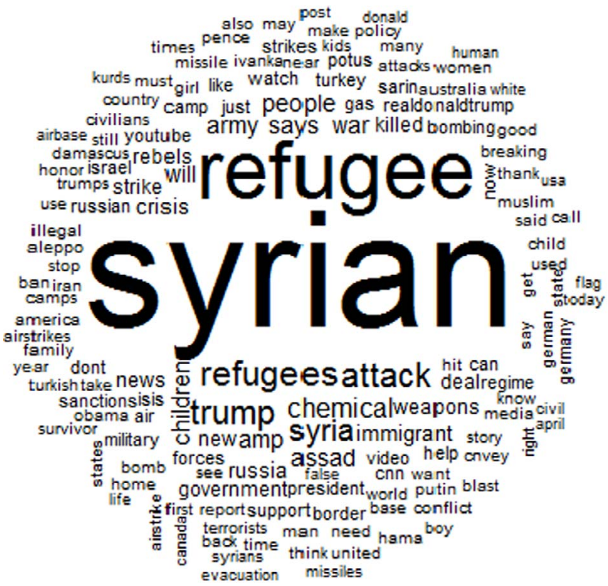

Fig. 9. Word cloud for English tweets.

3.4. Clustering tweets into categories

We also explored the clustering of the tweets into several categories and performed a comparison analysis of categories for English and Turkish tweets. The tweets were clustered into three main categories: Politics, War and Humanitarian.

To make fair comparison of categories between languages, we have chosen ten one-to-one translatable terms for each category from the frequently used terms in the tweets. Although it is very challenging to classify all terms into a category since some terms can be clustered into multiple categories in different contexts, we focused on a set of frequently used terms that are distinctively associated with the clusters. For example, the terms “weapon” and “bomb” remind us war and violence whereas the terms “children” and “women” are associated more with humanitarian side of the crisis. The list of terms that are used in the analysis is shown by category and language in Table 1. While we think the terms in each category are good representatives of the categories, we note that there could be additional terms that could be associated with the categories.

In the analysis, the tweets were clustered based on the frequencies of the terms in the categories. The ratios of categories are calculated using the formula below from the tweets that could be clustered.

$$\text{category\%} = \frac{\text{num. of tweets in category}}{\text{num.of tweets in all categories(politics + war + humanitarian)}}$$

Fig. 10 demonstrates the comparison of tweet categories between languages. Amongst the tweets that could be clustered, approximately 35% of tweets were clustered into Politics category and 33% were in War category and 32% of them were marked as Humanitarian in Turkish language. In the case of English language, 47% of clustered tweets were grouped into Politics category, 34% were clustered into War category and 19% were considered as Humanitarian respectively. Politics category carried the most tweets in both languages but the ratio of Political tweets was significantly higher in English than Turkish, 47% vs 35% respectively. On the

Table 1
List of selected terms in each category per language.

English Terms			Turkish Terms		
Politics	War	Humanity	Politics	War	Humanity
refugee	war	child	göçmen	savaş	çocuk
immigrant	terrorist	baby/infant	mülteci	terörist	bebek
border	bomb	women	sınır	bomba	kadın
politic	attack	people	politika	saldırı	insanlar
policy	kill/killer	camp	politik	katil	kamp
protest	army	peace	protesto	askeri	barış
racist	dead	humanitarian	ırkçı	ölü/şehit	insani
asylum	explosion	humanity	iltica	patlama	insanlık
assad	chemical	aid	esad	kimyasal	yardım
legal/illegal	weapon	elderly	yasal/yasadışı	silah	yaşlı

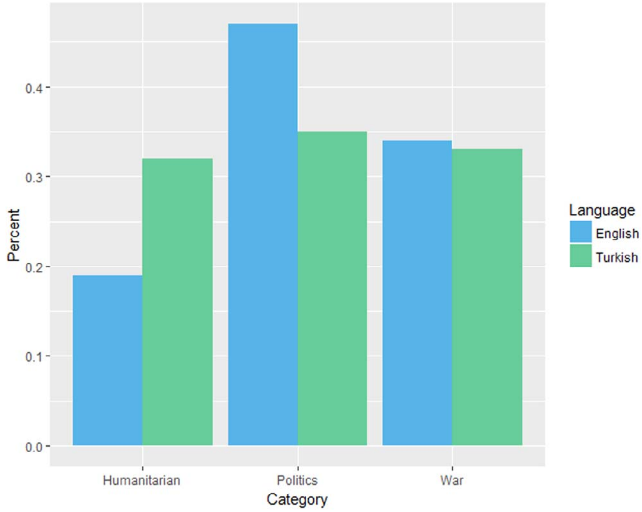

Fig. 10. Comparison of tweet clusters by language.

other hand, the proportion of Turkish tweets that were clustered into Humanitarian category with 32% was considerably higher than that of English with19%.

4. Discussion

In this study, Turkish and English tweets about the war in Syria and the refugee crisis were analyzed and consequently their sentiment results were obtained. The sentiment analysis of Turkish tweets was considered valuable since Turkey currently hosts the biggest number of Syrian refugees and Turkish tweets carried information to reflect public perception of a refugee hosting country. English was also chosen for analyses because it has the largest coverage and is the most commonly used language in the world. The sentiment analysis results based on the two languages were considerably different from each other. In Turkish tweets, the results of sentiment analysis indicated that the number of tweets classified within positive category was slightly more than the other two main sentiments categories of neutral and negative, yet the number of tweets within each of category was close to each other. None of the sentiment categories were significantly more than other. In English tweets, the number of neutral tweets was the largest. It was followed by the sentiment categories of negative and very negative, respectively. For English tweets, the proportional combination of negative and very negative tweets (40% of all tweets) was exceeding 3-fold of the proportion of positive and very positive combination (12%).

While Turkish speaking community appeared to share more positive tweets and almost equally distributed opinions among main sentiment categories, English speaking community shared distinctly more neutral and negative opinions compared to positive tweets. That being said, we note that the number of tweets was different for two languages. While we think that the difference in data size might not be a decisive factor for such a marginal difference in the proportions of sentiment categories, it might still have some impact on the distribution of sentiment categories as there were more English tweets than Turkish tweets in total in the analyses.

As revealed in the word cloud results, Turkish speaking community and English speaking community appeared to be focusing on different perspectives of the issue on Twitter. The world cloud results using the same search keywords of "Syrian" and "refugee" returned different set of frequently associated terms for the two communities. For Turkish speaking community, the details of the war happening near the borders of Turkey had much more importance while English speaking community seemed to discussing the legality of immigrants, policies and politics more frequently. However, people from communities of both languages shared opinions about refugee children and their concerns about the children of war. It appears that the children of the war are a major concern for people regardless the spoken language in our analysis.

In this study, a total of 2381,197 tweets were retrieved. However, we were able to use 348,707 tweets, 15% of the retrieved data, in the sentiment analyses after data cleaning. There were many duplicate tweets and retweets in the retrieved data, which were removed from the dataset as duplicates and retweets don't introduce a new opinion but would have affected the overall results. However, we did not limit the dataset to include the tweets from individuals only. Twitter news feeds were also considered for analyses as they provide valuable temporal information about the incidents. This might have slightly increased the number of neutral tweets.

Moreover, Emoji characters were removed from the tweets as a part of data cleaning in this study. The Emoji characters are more frequently being used in social media posts nowadays. They are known to impact the overall sentiment of Twitter posts (Shiha and Ayvaz, 2017). In future work, we plan to consider the Emoji characters in our sentiment analysis studies using social media data as the utilization of the Emoji characters might help obtain more accurate sentiment scores.

Although the time span of the tweets collected for this study was relatively short when considering 6 years passed since the beginning of the Syrian civil war and the sentiments of tweets may change over time, we think that the analyses results in this study are consistent with the current public perceptions of the two communities towards the Syrian war and refugee crisis since a large amount of randomly chosen Twitter posts were analyzed in order to avoid potential biases. Nevertheless, we plan to expand the study timeframe in the future to see whether or not the sentiments results are changing over time.

As a part of this study, we developed a Turkish sentiment lexicon. Since the lexicon was to cover the terms used in the retrieved tweets, it thus not only contained general vocabulary but included a lot of subject specific terms. Therefore, the lexicon can be considered as a case specific lexicon. We also assessed the effectiveness of the lexicon by manually evaluating some tweets and comparing the sentiment polarity and the scores against the sentiments generated by using the lexicon. The results of the analysis using the lexicon were consistent with manual observations. The lexicon establishes a firm basis for general sentiment analysis scoring Turkish sentences. It can be extended in the future to include additional terms from various fields to achieve better scoring in specific subject areas.

4.1. Potential limitations

In contrast to Turkish language, English is a global language and it cannot be constrained with a specific country. As a potential limitation, we note that the public perception of people towards Syrian civil war may vary in different English speaking countries. The reason for not analyzing the English tweets further by the geographic locations was due to the fact that the location information is not a required field in Twitter API. Thus, this information was not populated in Twitter and available to us for the analyses. We noticed that the location information is not reliably collected. Therefore, we did not consider location information in the analyses.

Furthermore, while Turkish and English tweets provide invaluable information regarding the public sentiments towards Syrian civil war, there exist other languages that people share opinions on social media regarding the topic. We think that the relevant tweets in languages such as Arabic and German could also potentially provide interesting insights. However, we currently do not have access to human resources that have expertise in Arabic or German languages and can help with development of lexicons. Thus, those languages were considered out of scope of this current study. In future work, we plan to expand this study to include the aforementioned languages.

In the timeframe analyses, we investigated important specific events that might be associated with sudden increases in the number of relevant tweets on certain dates. However, we note that there may be many factors that affect the opinions and perspectives of the people. Amongst other factors, these may include access to social media, policies of governments, news coverage, and coinciding local events at the time. While they are worth mentioning, these factors may not always be analyzable.

5. Conclusion

In this study, a series of sentiment analyses using Twitter data were performed in the subject of the Syrian civil war and following refugee crisis, which are currently among the most tragic and pressing issues in the world. We collected relevant tweets in two languages: Turkish and English. Upon a comprehensive twitter search, a total of 2381,297 tweets were collected for analysis. Out of all, 1353,367 of them were in English and 1027,930 them were in Turkish. After removing duplicates, retweets, and the tweets with missing information as a part of data cleaning, 250,857 English tweets and 97,850 Turkish tweets, and the total of 348,707 tweets were used in the analyses.

Upon sentiment analysis of retrieved tweets for each language, we observed that Turkish tweets were carrying more positive sentiments about Syrians and refugees when compared to English tweets with the ratio of 35% of all tweets versus only 12%, respectively. While the sentiments of Turkish tweets were almost evenly distributed among the positive, neutral and negative categories, the English tweets were largely composed of neutral and negative sentiments. Furthermore, we found out that the details of the war happening near the borders of Turkey attracted more attention of Turkish speaking community, whereas English speaking community argued the legality of immigrants, policies and politics more frequently. Nonetheless, both communities talked about refugee children and shared their concerns about the children.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.tele.2017.10.006.

References

Ahmed, S., Jaidka, K., Cho, J., 2016. The 2014 Indian elections on Twitter: a comparison of campaign strategies of political parties. Telematics Inform. 33 (4), 1071–1087.

Alarifi, A., Alsaleh, M., Al-Salman, A., 2016. Twitter turing test: identifying social machines. Inf. Sci. 372, 332–346.

Bianchi, C., Andrews, L., 2015. Investigating marketing managers' perspectives on social media in Chile. J. Business Res. 68 (2015), 2552–2559.

Bollen, Johan, Mao, Huina, Zeng, Xiaojun, 2011. Twitter mood predicts the stock market. J. Comput. Sci. 2 (1), 1–8.

Bose, S., Saha, U., Kar, D., Goswami, S., Nayak, A.K., Chakrabarti, S., 2017. RSentiment: a tool to extract meaningful insights from textual reviews. In: Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications Springer, Singapore, pp. 259–268.

Broniatowski, D.A., Paul, M.J., Dredze, M., 2013. National and local influenza surveillance through Twitter: an analysis of the 2012–2013 influenza epidemic. PLoS ONE 8 (12), e83672.

Carley, K.M., Malik, M., Landwehr, P.M., Pfeffer, J., Kowalchuck, M., 2016. Crowd sourcing disaster management: the complex nature of Twitter usage in Padang Indonesia. Saf. Sci. 90, 48–61.

Cheong, Marc, Lee, Vincent C.S., 2011. A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via Twitter. Inf. Syst. Front. 13 (1), 45–59.
Chianese, A., Piccialli, F., 2016. International workshop on Data Mining of Iot Systems (DaMIS): a service oriented framework for analysing social network activities. Procedia Comput. Sci. 98 (2016), 509–514.
Corley, C.D., Cook, D.J., Mikler, A.R., Singh, K.P., 2010. Text and structural data mining of influenza mentions in web and social media. Int. J. Environ. Res. Public Health 7 (2), 596–615.
Crannell, W.C., Clark, E., Jones, C., James, T.A., Moore, J., 2016. A pattern-matched twitter analysis of US cancer-patient sentiments. J. Surg. Res. 206 (2), 536–542.
Culotta, A., 2010. Towards detecting influenza epidemics by analyzing Twitter messages. In: Proceedings of the first workshop on social media analytics ACM, pp. 115–122.
Çetinkaya, C., Özceylan, E., Erbaş, M., Kabak, M., 2016. GIS-based fuzzy MCDA approach for siting refugee camp: A case study for southeastern Turkey. Int. J. Disaster Risk Reduction 18, 218–231.
Danneman, N., Heimann, R., 2014. Social Media Mining with R: Deploy Cutting-Edge Sentiment Analysis Techniques to Real-World Social Media Data Using R, www.it-ebooks.info.
Elkin, L.S., Topal, K., Bebek, G., 2017. Network based model of social media big data predicts contagious disease diffusion. Information Discovery and Delivery.
Fernández-Gavilanes, M., Álvarez-López, T., Juncal-Martínez, J., Costa-Montenegro, E., González-Castaño, F.J., 2016. Unsupervised method for sentiment analysis in online texts. Expert Syst. Appl. 58, 57–75.
Grant, A., O'Mahoney, H., 2016. Portrayal of waterpipe (shisha, hookah, nargile) smoking on Twitter: a qualitative exploration. Public Health 140, 128–135.
Kayser, V., Bierwisch, A., 2016. Using Twitter for foresight: an opportunity? Futures 84 (A), 50–63.
Khan, F.H., Qamar, U., Bashir, S., 2016. ESAP: a decision support framework for enhanced sentiment analysis and polarity classification. Inf. Sci. 367, 862–873.
Kim, J., Hastak, M., 2018. Social network analysis: characteristics of online social networks after a disaster. Int. J. Inf. Manage. 38 (1), 86–96.
Lampos, V., De Bie, T., Cristianini, N., 2010. Flu detector-tracking epidemics on Twitter. Mach. Learning Knowl. Discovery Databases 599–602.
Lee, N.Y., Kim, Y., Sang, Y., 2017. How do journalists leverage Twitter? Expressive and consumptive use of Twitter. Social Sci. J. 54 (2), 139–147.
Liao, C., Feng, C., Yang, S., Huang, H., 2016. Topic-related Chinese message sentiment analysis. Neurocomputing 210, 237–246.
Muhammad, A., Wiratunga, N., Lothian, R., 2016. Contextual sentiment analysis for social media genres. Knowl.-Based Syst. 108, 92–101.
Neppalli, V.K., Caragea, C., Squicciarini, A., Tapia, A., Stehle, S., 2017. Sentiment analysis during Hurricane Sandy in emergency response. Int. J. Disaster Risk Reduction 21 (2017), 213–222.
Oza, K.S., Naik, P.G., 2016. Prediction of online lectures popularity: a text mining approach. Procedia Comput. Sci. 92 (2016), 468–474.
Pang, B., Lee, L., 2008. Opinion mining and sentiment analysis. Found. and Trends Inf. Retrieval 1 (2), 1–135. Retrieved from. http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf.
Park, S.J., Park, J.Y., Lim, Y.S., Park, H.W., 2016. Expanding the presidential debate by tweeting: the 2012 presidential election debate in South Korea. Telematics Inform. 33 (2), 557–569.
Sasaki, Y., Kawai, D., Kitamura, S., 2016. Unfriend or Ignore Tweets?: a time series analysis on Japanese Twitter users suffering from information overload. Comput. Hum. Behav. 64, 914–922.
Shiha, M., Ayvaz, S., 2017. The effects of Emoji in sentiment analysis. Int. J. Comput. Electr. Eng. (IJCEE.) 9 (1), 360–369.
Soykoek, S., Mall, V., Nehring, I., Henningsen, P., Aberl, S., 2017. Post-traumatic stress disorder in Syrian children of a German refugee camp. Lancet 389 (10072), 903–904.
Statistic Brain, 2016. STATS | Twitter company statistics. In: STATISTIC BRAIN, [Online]. Available: http://www.statisticbrain.com/twitter-statistics/. Accessed: Jul. 30, 2017.
Sun, S., Luo, C., Chen, J., 2017. A review of natural language processing techniques for opinion mining systems. Inf. Fusion 36 (2017), 10–25.
Tellez, E.S., Miranda-Jimenez, S., Graff, M., Moctezuma, D., Siordia, O.S., Villasenor, E.A., 2017. A case study of Spanish text transformations for twitter sentiment analysis. Expert Syst. Appl. 81 (2017), 457–471.
Thompson, J.J., Leung, B.H.M., Blair, M.R., Taboada, M., 2017. Sentiment analysis of player chat messaging in the video game StarCraft 2: Extending a lexicon-based model. Knowl.-Based Syst. 2017, 1–14.
Vaz, E., Lee, K., Moonilal, V., Pereira, K., 2017. Potential of Geographic Information Systems for Refugee Crisis: Syrian Refugee Relocation in Urban Habitats. Habitat International.
Villasana, D., 2016. Picturing health: challenges for Syrian refugees in Turkey. Lancet 388 (10056), 2096–2103.
Wang, Hao, et al., 2012. A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In: Proceedings of the ACL 2012 System Demonstrations. Association for Computational Linguistics.
Wu, B., Shen, H., 2015. Analyzing and predicting news popularity on Twitter. Int. J. Inf. Manage. 35 (6), 702–711.
Yates, D., Paquette, S., 2011. Emergency knowledge management and social media technologies: a case study of the 2010 Haitian earthquake. Int. J. Inf. Manage. 31 (1), 6–13.