Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 05, No. 03, December 2019.

MULTI-LEVEL ANALYSIS OF POLITICAL SENTIMENTS USING TWITTER DATA: A CASE STUDY OF THE PALESTINIAN-ISRAELI CONFLICT

Iyad Al-Agha and Osama Abu-Dahrooj

(Received: 9-Jul.-2019, Revised: 15-Aug.-2019, Accepted: 3-Sep.-2019)

ABSTRACT

Given the tremendous growth of social media platforms, people have been actively spreading not only information in general, but also political opinions. Many research efforts have used social media content to analyse and predict the public opinion towards political events. This work presents an analytical study for measuring the political public opinion towards the Palestinian-Israeli conflict by using Twitter data. The study uses a novel data analysis model that leverages two levels of analysis: country-level analysis and individual-level analysis. The country-level analysis aims to explore the country's overall attitude towards Palestine by: 1) Identifying counties that generated the most topic-focused tweets, 2) Measuring the friendliness of each country towards Palestine. 3) Analysing the change of sentiment over time. The individual-level analysis aims to analyse data based on the activity and background of individuals. The attitudes of opinion leaders and ethnic groups were analysed and discussed in light of countries' attitudes.

The rich experience provided in this study through the proposed model for multi-level analysis, the step-by-step procedure, the variety of analysis techniques and the discussion of results can be informative for other developers and data analysts who are interested in analysing social media sentiment about political conflicts in particular.

KEYWORDS

Sentiment analysis, Public opinion, Twitter, Politics, Palestinian-Israeli conflict.

1. INTRODUCTION

With the dramatic rise of social media in the past decade, millions of people express their views on a great variety of topics. This has dramatically increased the data available to mine social media platforms, such as Twitter, for information about how people think and feel. Identifying the public opinion towards political issues is essential to shape the international policies, alliances and positions. Governmental officials should pay attention to public opinion to decide how to act. Opinion polls have been the standard mechanism for collecting public opinion. However, polls have several problems that have been reported in the literature [1]-[2]. They often use samples of small sizes or non-representative samples that may result in inaccurate findings. Unclear, biased or emotionally charged questions will produce misleading answers and weaken the accuracy of the results of a poll. In addition, the results of opinion polls are perspective as their findings apply only at the time the questions were asked. However, the public opinion towards a particular issue is likely to fluctuate over time based on recent updates. Finally, with polls it is difficult to perform fine-gained analysis or to understand the subjectivity and the motivations behind the public opinion. All the aforementioned limitations make opinion polls not very reliable and there is a need for other mechanisms to capture the public opinions.

Social media platforms have grown explosively over the past decade. People from all over the world have been using them extensively to express their views and discuss topics of interest. The large number of users, the variety of discussed topics and the massive volumes of posted content have made social media a rich source to understand and predict the population attitudes. Mining social media for political opinions may provide a faster and less expensive alternative to traditional polls.

Numerous research works have explored the mining of social media to analyse or predict political opinions [3]-[5]. However, these works were mostly event-specific and used techniques relevant only to the issue being investigated. In addition, most studies relied on sentiment analysis that primarily aims to predict the user feeling rather than the political opinion. In politics, judging a sentiment depends on

I. Al-Agha and O. Abu-Dahrooj are with Department of Computer Science, Faculty of Information Technology, The Islamic University of Gaza, Palestine. Emails: ialagha@iugaza.edu.ps and osama.abudhrooj@gmail.com

the side you stand by regardless of the user's emotions. Thus, the same idiom may be interpreted differently based on the context. For example, the idiom: "I feel sorry for the Palestinian people☹" conveys a feeling of sadness and sympathy, thus may be classified as "negative" by a conventional sentiment analyser, despite that it carries a positive attitude towards the Palestinian case. Similarly, expressions that evoke a positive emotion towards Israel, such as "I love Israel" should have a negative polarity from the perspective of Palestine. These examples show that the conventional sentiment analysis that is based on feelings or emotions may be inadequate for inferring political attitudes that are based on a specific understanding of what is "positive" and what is "negative". In addition, few studies [6]-[7] have used social media to explore the public opinion towards the Palestinian-Israeli conflict and the majority have used statistics to analyse existing situations rather than making predictions about the public attitudes.

In this paper, we propose an analytical study that uses a sample of Twitter text data in English to measure and analyse the political public opinion in several countries around the world towards the Palestinian-Israeli conflict. The proposed approach builds on a data analysis model that leverages two levels of analysis: country-level analysis and individual-level analysis. Several types of analysis are used under each level, each of which aims to provide an insight into a particular aspect. The aim is to provide more in-depth analysis that leads to a better understanding of the public opinion. This work demonstrates, through a realistic case study and a step-by-step procedure, how different data mining techniques, such as sentiment analysis, time-based analysis and opinion leaders analysis can be used to gain deeper and more fine-grained insights. We believe that the proposed analysis model can be adapted and reused for similar studies, especially those focusing on sentiment analysis of social media content about political conflicts.

2. Related Works

In the past decade, a growing number of studies have used data from Twitter to monitor sentiments for the purpose of tracking trends in politics, economy and public opinion [8]-[9]. In general, these studies can be classified into three research areas based on the purpose they used sentiment analysis for [1], [10]. The first research area concentrates on predicting real-world continuous values by analysing sentiments in social media, such as predicting stock market values. Several studies in this area have reported a notable improvement in prediction results when incorporating sentiments [11]-[13]. It is noteworthy that the focus in these studies was on emotive sentiment; i.e., mood states, rather than on polar sentiment (positivity, negativity) which is popular in politics.

The second area is result forecasting. A popular example is the predication of election results. Twitter has been used increasingly in the past decade to forecast the public opinion in the events of elections [3], [14]-[16]. Researchers in this area have focused either on determining current levels of support toward political actors or on predicting support in upcoming elections. However, they often had different opinions about the reliability of social media mining as a prediction tool in the time of elections. Some studies tried to validate sentiment measured from Twitter by comparing it with the public opinion measured from polls [4] or by comparing it with the final election results [1]. These studies reported high correlations between actual and predicted results and confirmed the potential of social media in predicting political views. In contrast, other studies underestimated the prediction based on social media due to many flaws, such as the untrustworthy content, the negligence of demographics, the non-representative sample and the inaccurate ways used to validate results [15], [17] . Overall, studies from both sides have supported the use of social media mining as a supplement for traditional polling.

The third related area is event monitoring, where the aim is to analyse the social media reactions to specific events. These events could be election debates [18]-[20], campaigns [3], [21] or any political event affecting the public opinion [7], [22]. Many works in this area focused merely on using sentiments of tweets to understand their relationships to the event of interest [3], [23]. Other works tried to enhance the sentiment analysis by combining it with other factors, such as the retweet behaviour [24], hashtags [25], the opinion leaders' behaviour [26]-[27] and contextual information, such as geo-location, temporal and author information [28]-[30]. In general, the previous studies concluded that Twitter proved to be an effective source of data for monitoring and assessing the public reaction associated with important events.

The work in this paper falls under the third area, as it aims to characterize the international public attitude towards the Palestinian-Israel conflict in terms of Twitter sentiment. It builds on the methods used in the literature and contributes in providing in-depth tracking of public opinion at multiple levels: the country level and the individual level.

When it comes to the Palestinian-Israeli issue, very few studies have exploited social media content to capture patterns or trends related to the ongoing conflict [6], [7], [31]. These studies, however, relied solely on systematic statistical reviews rather than on data mining or sentiment analysis. In addition, they were published in the domain of politics, where the focus was on the findings and implications rather than on the underlying technology. Although our findings may complement those of previous studies, the emphasis of this work is on describing the used approach and techniques.

Several efforts showed an interest in the online sentiment analysis to predict the result of elections and monitor political events. Tumasjan, Sprenger, Sandner and Welpe [2] presented the first attempt to investigate whether Twitter validly mirrors the German election results. They analysed about 100,000 political tweets identifying either a politician or a political party and used LIWC2007 [32] tool for extracting sentiment from the tweets. Burnap, Gibson, Sloan, Southern and Williams [33] used Twitter to forecast the outcome of 2015 UK general elections. They used an approach that incorporates sentiment analysis and prior party support to generate a forecast of parliament seat allocation. Ceron, Curini and Iacus [34] and Ceron, Curini and Iacus [35] attempted to analyse several elections in Italy, France and the US and showed that a supervised learning method developed by Hopkins and King [36] does a good job of explaining fluctuation in party or candidate support in various contexts. However, the previous works put emphasis on using emotional states to identify user preferences and did not examine the influence of other factors, such as individual characteristics, geo-locations and opinion leaders.

Besides the sentiment analysis of Twitter data, some efforts tried to incorporate other sources of user-generated media. For example, Le, Boynton, Mejova, Shafiq and Srinivasan [37] studied Twitter communications around the 2016 U.S. elections and implemented computational methods for tracking political discourse about party, personality traits and policy on Twitter. O'Connor, Balasubramanyan, Routledge and Smith [4] investigated the people’s remark measured from polls with opinion measured from microblogging sites. They used time series to assess the population's aggregate opinion on a topic and measured correlations to several polls conducted during the same period of time. Marozzo and Bessi [38] presented a study analysing the polarization of social network users and news sites during political campaigns characterized by the rivalry of different factions. They performed temporal analysis to monitor the changes of polarization during the weeks preceding the vote. Martin-Gutierrez, Losada and Benito [39] analysed temporal series and interaction networks corresponding to two Twitter datasets downloaded during the Spanish electoral campaigns of 2015 and 2016 in order to identify recurrent patterns of user behaviour. Cody et al. [40] studied the sentiment surrounding climate change conversation on Twitter and used temporal analysis to observe how sentiment varies in response to climate change news and events. Our work also extends the sentiment analysis by using temporal analysis and incorporating multiple factors. It also analyses the influence of each factor on the overall public opinion. However, it has a different objective as it focuses on the Palestine-Israeli issue.

The contributions of this work reside in the following: First, it proposes a multi-level model of analysis that leverages multiple factors at both group and individual levels and employs computational methods to perform fine-grained analysis of public opinion. This is different from most of the existing efforts that focused solely on sentiment analysis or used a certain type of features. Second, the work discusses the special considerations for political sentiment analysis and demonstrates, through an experimental study, that sentiment analyses that are based on emotional states may be inadequate for inferring political polarization. We believe that the detailed procedure and computational methods reported in this work can be informative to data analysts and practitioners in investigating political conflicts in particular.

3. Overview of The Proposed Approach

The overall approach used to undertake this study is depicted in Figure 1. It consists of the following steps: data collection, data pre-processing, political sentiment analysis and feature extraction and analysis. The following sections start by describing the data analysis model consisting of two levels of analysis: country-level analysis and individual-level analysis. The features that should be extracted to realize the proposed model are also described.

Afterwards, a case study that utilizes the proposed model to analyse the international public opinion towards the Palestinian-Israeli conflict is presented in detail. Data collection and pre-processing steps are described, highlighting the considerations we took to assure that the collected data will not lead to invalid or biased results. At the core of our approach resides the sentiment analysis step. The paper reports on the experiments conducted to compare several sentiment classifiers and to train and evaluate our own sentiment classifier.

The sentiment analysis step is followed by feature extraction, in which several features are extracted or derived from the inferred sentiments. Extracted features are then used to carry out the analysis at both country and individual levels. Finally, results are presented and discussed.

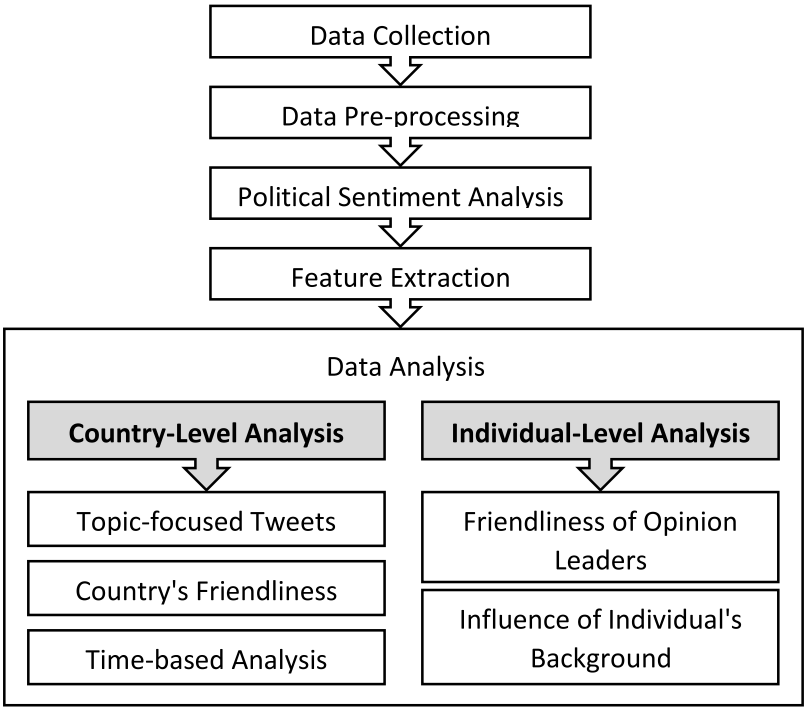

Figure 1. Approach for analysing public opinion towards the Palestinian-Israeli conflict.

4. DATA ANALYSIS MODEL

The data analysis model consists of two levels of analysis as shown in Figure 1. These levels are explained as follows:

4.1 Country-level Analysis

The purpose of the country-level analysis is to explore the country's overall interest in and attitude towards the political issue being studied. Country-level analysis is done through the following:

- Identifying counties that generated the most tweets related to the political issue; i.e., topic-focused tweets. The aim in our case study is to determine countries that show, on Twitter, the most concern about and awareness of the Palestinian-Israeli issue regardless of sentiment. Thus, tweets are counted per country, while sentiment scores are ignored.
- Measuring the friendliness of each country towards Palestine: Friendliness of a country indicates the level of support and sympathy it shows towards one side and can be determined from the polarities of tweets. In this study, friendliness is defined from the perspective of Palestine, so that positive and supportive attitudes towards Palestine lead to high friendliness rates. In contrast, views opposing the Palestinian side or advocating for the opposite side yield low friendliness rates.
- Analysing data across time to investigate how the public opinion changes over time.

4.2 Individual-level Analysis

The individual-level analysis aims to analyse data based on the activity of individuals and their backgrounds. This is performed through the following:

- Capturing the attitudes of opinion leaders. The term "opinion leader" refers to an active user on social media who has a large number of followers and can influence the opinions and behaviours of others [41]-[42]. Identifying opinion leaders is crucial to promote behaviour change or to identify subjects that are of high interest to people [43]. Measuring the attitude of opinion leaders towards political issues is important, because they reflect large sectors in their communities.
- Capturing the individual's characteristics: The individual's background or characteristics, such as nationality, religion, ethnicity and gender, may influence his/her political stance. For example, women are likely to stand in favour of issues pertaining to women rights and Arabs and Muslims are more keen to support the Palestinian rights. Identifying these characteristics from social media, where possible, will help better understand the motivations behind the public opinion. However, deciding which characteristics to capture and analyse is case-specific and depends on the objectives of the case study. For example, this work sought to measure the influence of the individual's ethnicity on the public opinion and the potential relationship between the ethnicity of users and their perceptions of Palestine.

5. Extracted Features

To achieve the data analysis model as explained above, several features need to be extracted from the collected tweets. These features are as follows:

- Polarity**:** Polarity is the sentiment score of the tweet, which determines the classification of the tweet (e.g. positive, negative or neutral). Polarities of tweets are measured by using the sentiment analyser. Other features will be derived from the polarities of tweets.
- Friendliness**:** The friendliness of a country is measured by calculating the average polarity of tweets posted by users in the country. Similarly, friendliness of an individual is the average polarity of tweets posted by the individual. To compute the average polarity, we interpreted the three sentiment values: positive, neutral and negative into +1, 0 and -1, respectively. Then, the friendliness for a country F_c is computed using the following equation:

$$\boldsymbol{F_c = \frac{\sum Polarity(t_i)}{n} \times 100} \tag{1}$$

  where, n is the number of tweets attributed to the country c. t_i is a tweet posted from the country c. The friendliness score ranges between -100 and +100, where +100 denotes the maximum friendliness value.
- Leadership**:** This feature is used to identify opinion leaders. Different metrics have been used in the literature to identify opinion leaders [42], [44]. In this work, Twitter users in each country who have the most number of followers are treated as opinion leaders.
- Individual's characteristics**:** In this study, the aim is to identify the individual's ethnicity from the user's name or nickname and then to analyse the influence of inferred ethnicities on the public opinion.

6. Data Collection

Twitter's public API is a streaming API offered by Twitter for collecting tweets. Although it has been widely used in the literature, it has a drawback in that it provides only 1% or less of its entire traffic, without control over the sampling procedure, which is likely insufficient for accurate analysis of public sentiment [45]. Instead, we used a Twitter search analytics and business intelligence tool called Followthehashtag [46]. Followthehashtag enables searching for tweets over a specific period of time. We first used Google Trends(i) to find top search keywords used in Palestine and Israel that are related to the Palestinian-Israeli conflict over the year 2016. Then, we selected keywords that represent the opposite views of the two sides of the conflict in order to avoid biased results. Examples of selected keywords include: Palestinian-Israeli conflict, Israeli occupation, Apartheid wall, settlements, Gaza, West bank, Judea and Samaria, Jerusalem, Palestinian terrorism and suicide bombings. Finally, we used these keywords to perform a query-based search to collect tweets related to the conflict that were posted during the year 2016.

In total, 178,524 tweets were collected. These tweets were posted by approximately 48,531 users during the period from Dec. 20 2015 to Dec. 31 2016. We think a period of one year is sufficient to explore the political trends on Twitter and to perform time-based analysis, since many related studies used equal or shorter periods (e.g. [1, 2, 5, 31]).

The following information was retrieved for each tweet: the tweet's text, the username and nickname of Twitter user, date and time of posting , country and place of origin, number of followers of the tweet's author, number of users followed by the author and hashtags. Most of these tweets were from the US, UK, Canada, Australia, Finland and some other European countries. 89.78% of the collected tweets were in English. Table 1 shows statistics about the collected tweets. The whole dataset can be found on the following link: https://github.com/odahroug2010/2017.

Table 1. Statistics about collected tweets.

General information	Total# of tweets	178,524
	Number of users	48,530
	Duration	Dec. 20 2015 to Dec. 31 2016
	English tweets	89.78%
	Retweets	7948
	Avg. no. of words per tweet	12.74
	Standard Dev. of words	5.002
Location information	No. of countries	174
	Top sources of tweets	US, UK, Canada, Australia, Finland and other European countries
	No. of tweets with unknown sources	28156
	No. of retweets	7948
	Min. tweets by country	24
	Max. tweets by country	27490
	Avg. tweets by country	777.88
	Standard Deviation	3363.68

7. Data Pre-processing

Tweets often have special characteristics that make their pre-processing different from that of ordinary texts. Tweets are of limited length (140 characters at most) and may contain special texts, such as hashtags, URLs, emoticons and usernames. For the pre-processing of tweets, we used the approach depicted in Figure 2, which consists of the following steps:

- Filtering: Collected tweets were filtered by: 1) removing non-English tweets: 10.22% of collected tweets were written in non-English languages and thus were excluded, 2) removing tweets with unknown resources: 28156 tweets in total did not have countries of origin. These tweets were excluded, because they are out of the scope of our analysis, 3) removing re-tweets: 7948 of tweets were retweeted and these were excluded from the dataset, so that only original tweets are counted. 124,174 tweets remained after the filtering step.

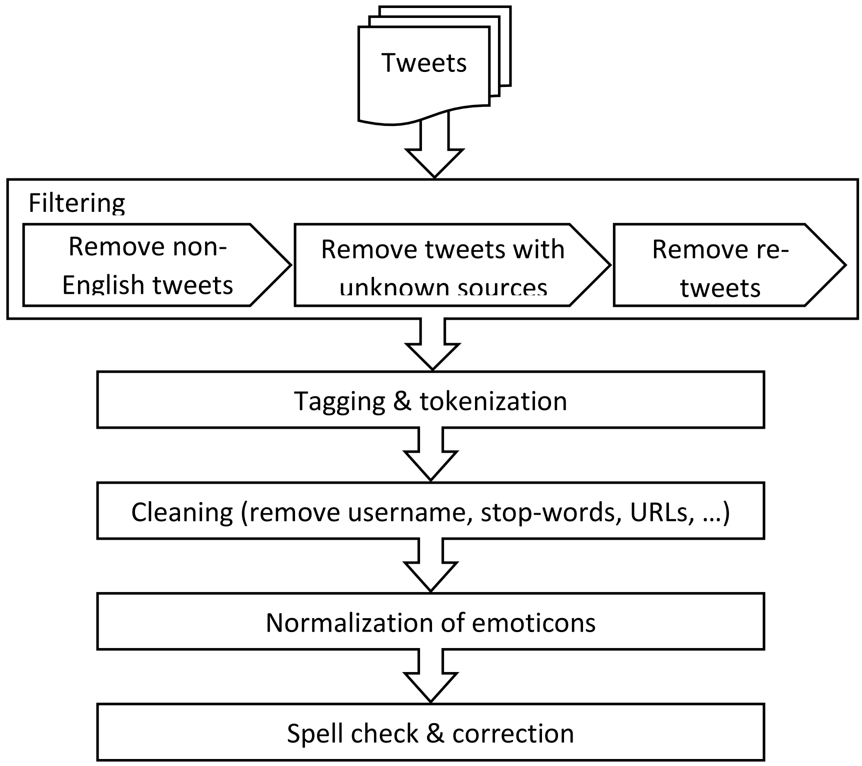

Figure 2. Pre-processing steps of tweets.

- Tokenization and tagging**:** As Twitter allows users to write short texts only, tweets often come with a special grammar and abbreviations, so that users can convey the messages with least possible words. Traditional tokenizers and POS taggers may be inadequate for pre-processing tweets and there is a need for alternatives that can recognize tweet's tokens, hashtags, emoticons and URLs. We used a text processing library called ArkTweetNLP to tokenize and tag tweets [47]. The library was developed specifically to handle informal and online conversational text including various non-standard lexical items and syntactic patterns.
- Cleaning: Twitter users prefer to use symbols and non-standard language in their tweets. Many of the used symbols may be irrelevant and thus should be excluded to avoid an incorrect result when applying the sentiment analyser. In our approach, tweets were cleaned by removing the following parts: usernames, numeric expressions, punctuations, URLs and stop-words that are unlikely to affect sentiments. These parts were recognized from the tagger applied in the previous step.
- Normalization of emoticons: Emoticons are important for sentiment analysis; thus, their meanings should be preserved and they should not be removed from the tweets. In our approach, we used a special dictionary that contains the most used emoticons and their meanings in English [48]. This dictionary was used to replace each emoticon with its relevant meaning. These examples show that conventional sentiment analysis that is based on feelings or emotions may be inadequate for inferring political attitudes that are based on a specific understanding of what is "positive" and what is "negative".
- Spell check and correction: Tweets may contain incorrect or miss-spelled words and this will affect the result of sentiment analysis. This step manipulates these words by using a spell checker and substitutes them with correct words. Jazzy Spell Checker [49] was used for this step. As an example of the output of this step, a tweet like "I looove palestin. Happi to visit it" will be corrected to "I love Palestine. Happy to visit it".

8. SENTIMENT ANALYSIS

Sentiment analysis is the core step to identify attitudes towards the Palestinian-Israeli conflict. When sentiments are identified, tweets can be categorized based on different features. Therefore, the results of subsequent steps largely depend on the quality of sentiment analysis. It is assumed that the tweet is an opinion and therefore we need to know its polarity classification, which is positive, negative or neutral. To achieve this, we used a supervised approach for sentiment analysis.

As mentioned earlier, it is important to emphasize that the sentiment analysis in this work aims to identify the political stance rather than mere the user feeling. In political conflicts, as in our case study, the polarity of a tweet should be determined based on the side you stand by regardless of the expressed emotions. As an example, tweets that show support to Israel are assessed as 'negative' from the perspective of Palestine even if they convey positive emotions. Although there are several pre-trained "off-the-shield" tools to perform sentiment analysis, these tools are often trained to identify feelings or emotions rather than political sentiment and thus they may be inadequate for the purpose of this study. Therefore, we decided to build our own sentiment classifier by training it on a manually-labelled dataset. Then, the performance of the classifier will be assessed by comparing it with other pre-trained sentiment analysers.

Since the collected tweets do not come with predefined sentiments, we decided to pick a sample of tweets and label them manually with the relevant polarity (positive, negative or neutral). These labelled tweets will be then used to train and evaluate the sentiment analyser. 1300 tweets (about 10% of the entire dataset after the filtering step) were chosen randomly and given to two human subjects to label them separately. In general, the labelling of tweets was done from the perspective of Palestine based on the following criteria:

- Tweets that include appreciation, praise, glorification or support for Palestine or the Palestinian issue were labelled as positive. For example, idioms like "Free Palestine" or "It is called Palestine, not Israel" are assigned positive polarity.
- Tweets that show solidarity and sympathy with Palestine or Palestinians were labelled as positive. For example, idioms like "Please donate for the children of Gaza" or "Save Palestinian children …" should be labelled as positive.
- Tweets that contain idioms denoting negative attitude towards "Israel", e.g. "Stop the Israeli apartheid wall" are considered positive from the perspective of the pro-Palestinian point of view.
- Tweets that show clear support for or sympathy with "Israel" were labelled as negative. For example, idioms like "I love Israel" or "Israel has the right to defend itself" all carry positive attitude towards "Israel" and negative attitude towards Palestine and thus were labelled as negative.
- Tweets that use Israeli naming conventions, such as "Judea and Samaria", "IDF army" and "Palestinian terrorists" were treated as negative sentiments, since they adopt a pro-Israel stance.

After analysing labels received from the two subjects and ignoring disagreements, we ended up with 1264 tweets, of which 637 were positive, 543 were negative and 84 were neutral.

Sentiment analysis in this work was carried out using a logistic regression model implemented by LingPipe [50]. LingPipe classifies texts by using a language model on character sequences and the execution uses the 8-gram language model. The labelled 1264 tweets were randomly split into two parts: 80% of the tweets (1011 tweets) were used for training and 20% (253 tweets) were used for testing. 10-fold cross validation was performed.

Table 2 shows the testing results of the trained classifier. Precision and recall values for each class were calculated by creating the confusion matrix. The matrix shows that the classifier achieved good results with positive and negative tweets, but the performance was low with neutral tweets. However, the low performance in case of neutral tweets will have a marginal impact on the results due to the low number of neutral tweets in general.

The performance of our sentiment classifier was also evaluated by comparing it with other pre-trained sentiment classifiers that are: Stanford CoreNLP [51], SentiStrength [52] and the pre-trained LingPipe. These classifiers were chosen, because they are frequently used to analyse the user sentiments on social media, especially in political events [53-58]. The testing dataset used above for testing our classifier was also used for testing the other classifiers.

The comparison results are shown in Table 3. Results indicate that our classifier outperformed the other classifiers significantly. It is also obvious that the performance of the pre-trained classifiers was remark-

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 05, No. 03, December 2019.

Table 2. The confusion matrix of the trained sentiment classifier.

	Label Positive	Label Negative	Label Neutral	Total Predicted	Precision	Recall
Predict Positive	104	12	2	118	88.1%	80%
Predict Negative	18	93	3	114	81. 6%	83.8%
Predict Neutral	8	6	7	21	33.3%	58.3%
Total Label Class	130	111	12			

ably poor. This can be attributed to the fact that they are designed to infer polarity based on emotional states that often contradict with political attitudes. This proves that the traditional sentiment analysis may be inadequate for inferring political polarization, where the polarity becomes a relative issue depending on the perspective of the interpreter and the case being analysed.

Table 3. Comparison between sentiment classifiers.

S. No.	Classifier	Accuracy	Precision	Recall	F-measure
1	Stanford CoreNLP	8.1%	30.6%	22.6%	26%
2	SentiStrength	7.9%	42.2%	27.8%	33.5%
3	Pre-trained LingPipe	31.2%	35.6%	30.5%	32.9%
4	**Our classifier**	**80.63%**	**81.35%**	**80.61%**	**81%**

The sentiment classifier built was used to measure polarities of 169694 tweets; these are the whole collected tweets excluding the retweets and the tweets used for training and testing of the sentiment classifier.

9. DATA ANALYSIS AND RESULTS

The sentiment classifier built in the previous section was used to measure the sentiments of all tweets in the dataset, excluding those used to build and test the classifier. In total, the sentiments of 122,921 tweets were measured.

The following sub-sections describe the application of the proposed analysis model, see Figure 1, on measured sentiments in order to derive the features needed to analyse the international public opinion. Afterwards, the main findings are presented, discussed and validated where possible. Apache Spark [59], which is an analytics engine for large-scale data processing, was used to implement the analysis model.

9.1 Country-level Analysis

The country-level analysis includes three types of analysis: countries that generated the most topic-focused tweets, friendliness of countries and time-based analysis. Results of each analysis is explained as follows.

9.1.1 Countries that Generated the Most Topic-focused Tweets

The volume of tweets that can be attributed to each country was measured. At this stage, polarities of tweets were ignored and the focus was only on counting the number of tweets per country.

Each tweet in the dataset often comes with geo-information that help identify its country of origin. One attribute is called "country" and it should be set with the country code. For example, tweets posted from the UK have the country value "GB". However, the country code may be missing for many tweets and it can be identified only if it is set in the user profile. Tweets can also have geocoding attributes named "Latitude" and "Longitude". These attributes are set to valid values for tweets posted from devices with enabled GPS service. For tweets that have latitude-longitude values but the country value is missing, the Google maps geocoding service was used to determine the corresponding countries. After assigning tweets to countries, tweets were counted per country and countries that ended up with a number of tweets less than 0.1% of the total number of tweets were ignored.

Table 4 shows the top ten countries in terms of the number of tweets concerning the Palestinian-Israeli issue. Canada, the UK and the US generated the most tweets. This result is expected considering the high involvement of these countries in the Middle East affairs. The bottom countries were Slovenia, New Zealand and Austria. Figure 3 illustrates the results on a geographical map. When considering the number of population, Jersey, Canada and Finland generated the most tweets *per capita*. The bottom countries were Nigeria, India and China.

Table 4. Top ten countries in terms of the number of tweets related to the Palestinian-Israeli conflict.

No.	Country	Country code	Focused Tweets
1	Canada	CA	27,490
2	United Kingdom	GB	23,010
3	United States	US	20,125
5	Ecuador	EC	9,342
6	Finland	FI	3,654
7	Australia	AU	3,125
8	Netherlands	NL	2,646
9	India	IN	1,445
10	France	FR	1,215

Figure 3. Number of tweets per country.

To get insight into the validity of the above results, we compared these results with the corresponding country indices generated from Google Trends. Google Trends provides a metric called Google index that indicates the frequency at which people in a country search for the term during a specific period of time. We used Google index to measure the frequency at which people search for the terms related to the Palestinian-Israeli issue in each country from the top 30 countries that posted the most tweets according to our results. The search activity was measured from January 2016 to December 2016, which is the same period through which the tweets were collected. Our assumption is that the search activity,

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 05, No. 03, December 2019.

measured through Google index, should be consistent with the tweeting activity during the same period of time. To make easy comparison, the tweet counts were normalized by log-transform, so that they become comparable with values from Google index (The log of tweet count is named as Twitter index) [60]. We plot the Google index as the x-axis and the Twitter index as the y-axis. The result is depicted in Figure 4.

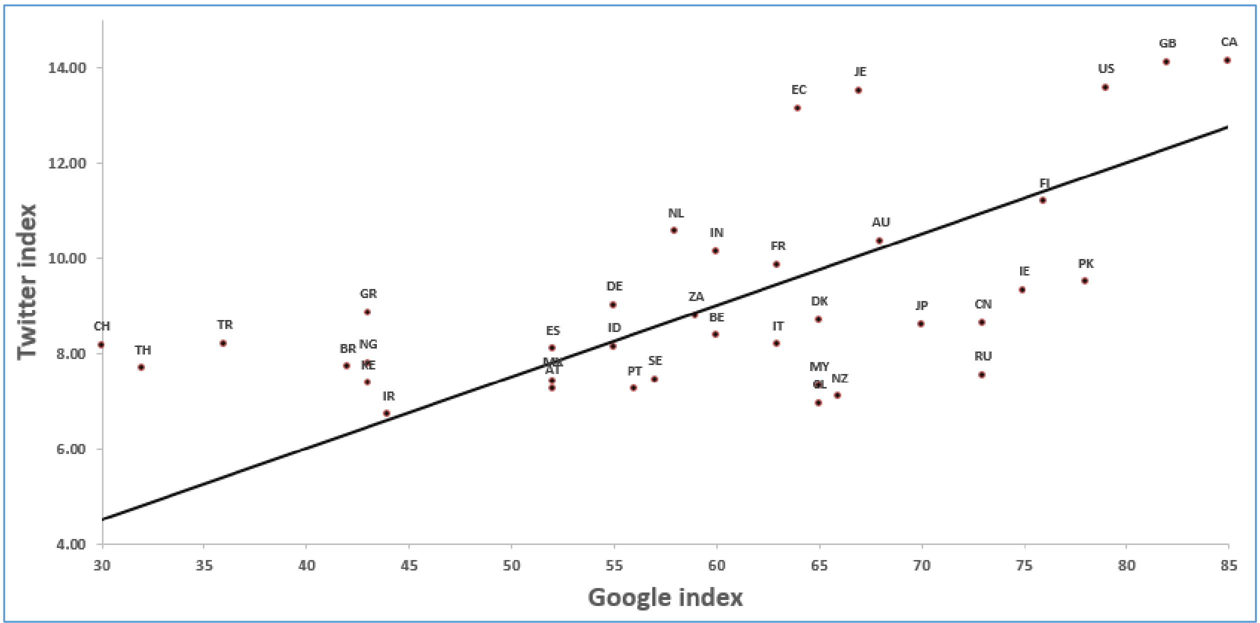

Figure 4. Correlation between Twitter index with Google index.

We then measured the coefficient of correlation between the Google index value and the Twitter index. The result was 0.685, which indicates a strong correlation [61].

9.1.2 Friendliness of Countries

The friendliness of a country is calculated by using Equation 1. It is the average sentiment score for each country. Tables 5 and 6 show information about the most and least friendly countries; respectively, along with tweets statistics. Figure 5 plots the friendliness scores for the top twenty countries. Table 6 lists the least friendly countries.

The top friendly countries were Finland, Brazil and Thailand. The least friendly countries were Switzerland, Austria and Russia. Of the top twenty countries, Figure 5 shows that only five countries have friendliness scores over zero, while the rest have below-zero scores. This result indicates that the public opinion is still highly negative towards Palestine even in the top friendly countries. Several countries like France, Greece, Nigeria and Italy got close to zero friendliness scores.

Referring to the distribution of sentiments and the standard deviation in Tables 5 and 6, a high divergence of attitudes can be observed in most countries. For countries like France, Italy and the UK, the numbers of positive and negative tweets were mostly comparable, while neutral tweets were much fewer in numbers. This result shows that the public opinion in these countries is highly divided. The small number of neutral voices also indicates the large polarization in the public opinion towards the Palestinian issue.

9.1.3 Time-based Analysis

The motivation of time-based analysis is to explore how the public opinion varies over time. Each tweet in our dataset is associated with a timestamp that specifies when the tweet was posted. Therefore, tweets can be treated as time series that can be analysed to extract meaningful patterns.

Due to the variations among countries, utilizing the whole volume of tweets for time-based analysis can result in a large variance. Therefore, time-based analysis was carried out only for the top three countries in terms of the number of posted tweets. These countries are Canada, the UK and the US.

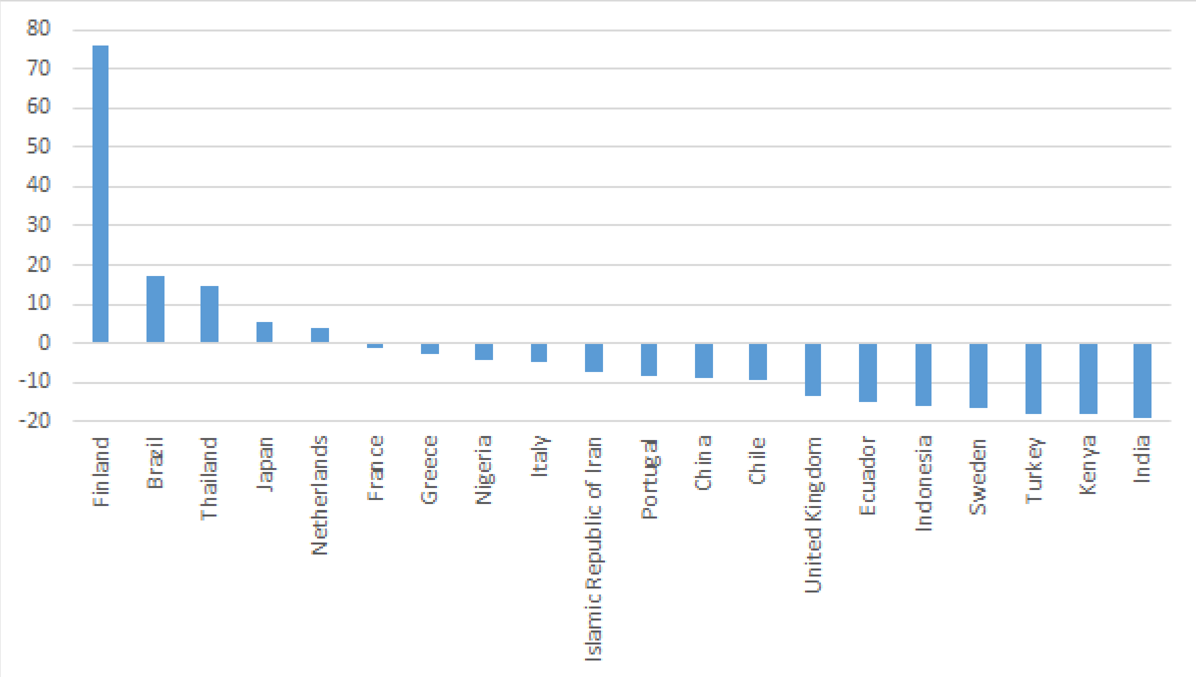

Figure 5. Friendliness scores of top twenty friendly countries.

Table 5. Top ten countries in terms of friendliness.

No.	Country	Focused Tweets	Positive	Negative	Neutral	**Friendliness**	St. Dev.
1	Finland	3,654	3,177	401	76	**75.97**	0.63
2	Brazil	382	184	118	80	**17.28**	0.87
3	Thailand	262	127	89	46	**14.50**	0.90
4	Japan	642	308	272	62	**5.61**	0.95
5	Netherlands	2,646	1,182	1,081	383	**3.82**	0.92
6	France	1,215	440	457	318	**-1.40**	0.86
7	Greece	820	317	338	165	**-2.56**	0.89
8	Nigeria	315	104	118	93	**-4.44**	0.84
9	Italy	577	207	235	135	**-4.85**	0.87
10	Islamic Republic of Iran	218	80	96	42	**-7.34**	0.90

Figure 6 shows how the friendliness scores of these countries have changed over the year 2016. It is obvious that the public opinion in the three countries fluctuated over time and the pattern of change was similar for the three countries. Friendliness scores were low in the first half of the year, before rising up to a peak value in June-July. Attitudes then went down again, then went up at the end of September, before going down again.

To understand these results, we tried to link the time-based changes with the significant events that took place over the year and that were related to the Palestinian-Israeli issue. These events could be discovered easily by searching the news archives on the Web. The declining attitude in the first quarter of 2016 may be explained by the stabbing spree that took place in Jerusalem and other Palestinian cities. A total number of 354 tweets related to the stabbing incidents were tweeted in the three countries during

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 05, No. 03, December 2019.

Table 6. Bottom ten countries in terms of friendliness.

No.	Country	Focused Tweets	Positive	Negative	Neutral	**Friendliness**	St. Dev.
1	Switzerland	381	72	258	51	**-48.82**	0.79
2	Australia	3,125	754	1,862	508	**-35.46**	0.84
3	United States	20,125	4,762	10,203	5,160	**-27.75**	0.82
4	South Africa	717	171	370	176	**-27.04**	0.82
5	Russian Federation	257	60	125	72	**-25.29**	0.81
6	New Zealand	177	44	88	45	**-24.86**	0.83
7	Belgium	399	90	186	123	**-24.06**	0.80
8	Mexico	278	65	131	82	**-23.74**	0.81
9	Germany	830	225	416	189	**-23.01**	0.85
10	Denmark	639	177	322	140	**-22.69**	0.85

the first quarter of 2016. The rising attitudes towards Palestine in June-July 2016 may be attributed to the demolitions of Palestinian houses that took place in July 2016 and resulted in the displacement of dozens of Palestinians(ii). In addition, the press releases that accused Israel of forcing Palestinians to withstand cruel and inhuman conditions at its borders have also grabbed attention during June 2016(iii)(iv). In total, 388 tweets were posted in response to the former events in June-July 2016.

Another rise of attitude towards Palestine was observed in September 2016 that can be explained by the reaction over the death of Shimon Peres, the former Israeli Prime Minister who is seen as a war criminal by pro-Palestinians(v). In total, 571 tweets referring to "Shimon Perez" were tweeted from these countries during September-October 2016, most of which had positive polarity with respect to Palestine. In addition, the UNESCO resolution on 12th October 2016 that condemned the Israeli policies around Al-Aqsa Mosque compound also got considerable attention in social media(vi)(vii). 332 tweets related to the UNESCO resolution were tweeted from the three countries during October-November 2016.

9.2 Individual-level Analysis

Individual-level analysis aims to explore the attitudes of specific types of individuals. Two groups of individuals will be identified: opinion leaders and individuals with certain ethnicities.

9.2.1 Influence of Opinion Leaders

Different metrics haven been used in the literature to identify opinion leaders on social networks. Some of these metrics have utilized the number of followers, interactions and activity, the leadership or social network analysis [44, 62, 63]. In this work, opinion leaders will be identified by using the number of

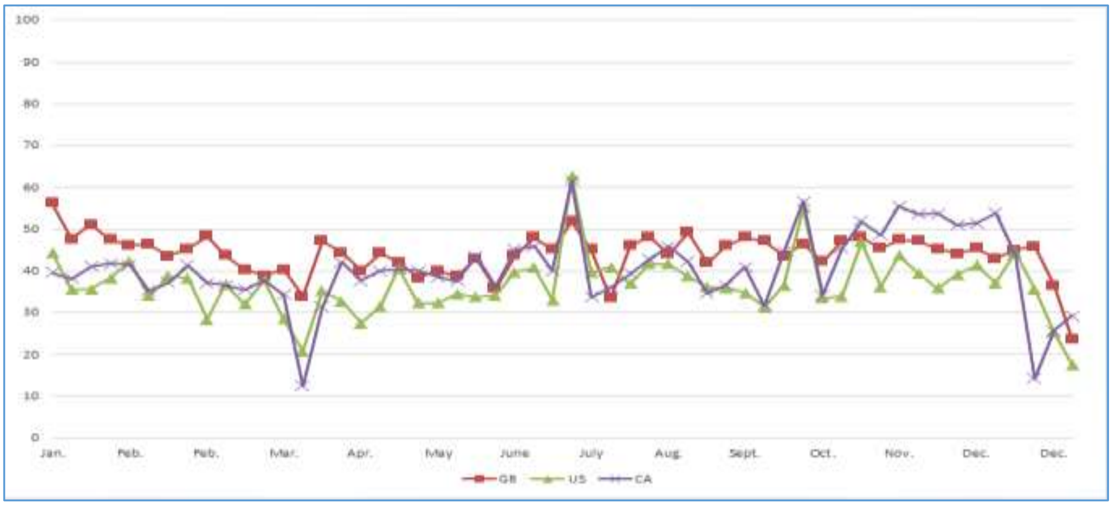

Figure 6. Time-based analysis of public opinion in three countries (UK, US and Canada).

followers, so that users with the largest number of followers in each country will be treated as opinion leaders.

We used the method proposed by Moore and McCabe [64] to identify users with extreme number of followers in each country. Moore and McCabe's method has been widely used in data analysis to find outliers in a distribution, whereas an outlier is the number that is more than 1.5 times the length of the box away from either the lower or upper quartiles. In our approach, opinion leaders are Twitter users whose numbers of followers are considered as "outliers above the upper quartiles" based on the Moore and McCabe's method.

From a total of 38,328 users, 1,794 users were identified as opinion leaders. Table 7 shows statistics about the opinion leaders, while Table 8 shows the top ten countries in terms of the number of opinion leaders. The US, Canada and the UK have the majority of opinion leaders; i.e., 59.14%.

Table 7. Statistics of opinion leaders.

Avg. no. of followers per opinion leader	203623.49
St. dev. of followers per opinion leader	89015.57
Avg. no. of tweets per opinion leader	13.76

Table 8. Top 10 countries in terms of number of opinion leaders.

No.	Country	No. of opinion leaders
1	United States	425
2	Canada	391
3	United Kingdom	299
4	France	95
5	India	61
6	Pakistan	35
7	Finland	31
8	Australia	29
9	Netherlands	23
10	South Africa	22

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 05, No. 03, December 2019.

Identified leaders were mostly official organizations, such as newspapers, government officials or media personnel. For example, among the top opinion leaders in the US were Reuters, Bernie Sanders and Billboard, while among the top opinion leaders in the UK were The Economist, ABC News and United Nations.

After identifying opinion leaders, the friendliness scores for them were calculated by using Equation 1. Then, the average friendliness score of opinion leaders in each country was calculated. The standard deviation per country was also calculated to identify the variance in friendliness of opinion leaders.

Figure 7 shows the results for the top twenty countries in terms of friendliness of opinion leaders, while Figure 8 shows the standard deviation values for friendliness of opinion leaders.

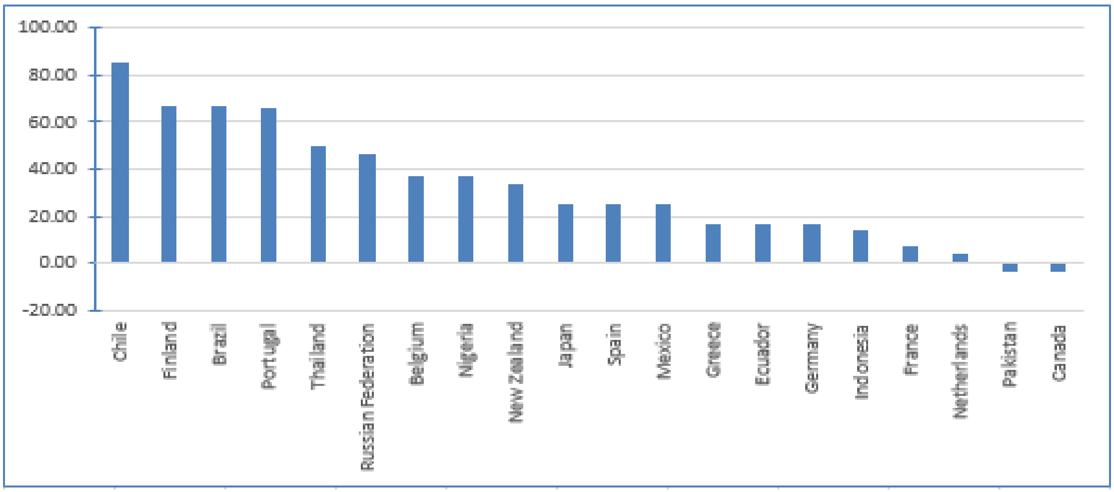

Figure 7. Average friendliness scores of opinion leaders per country.

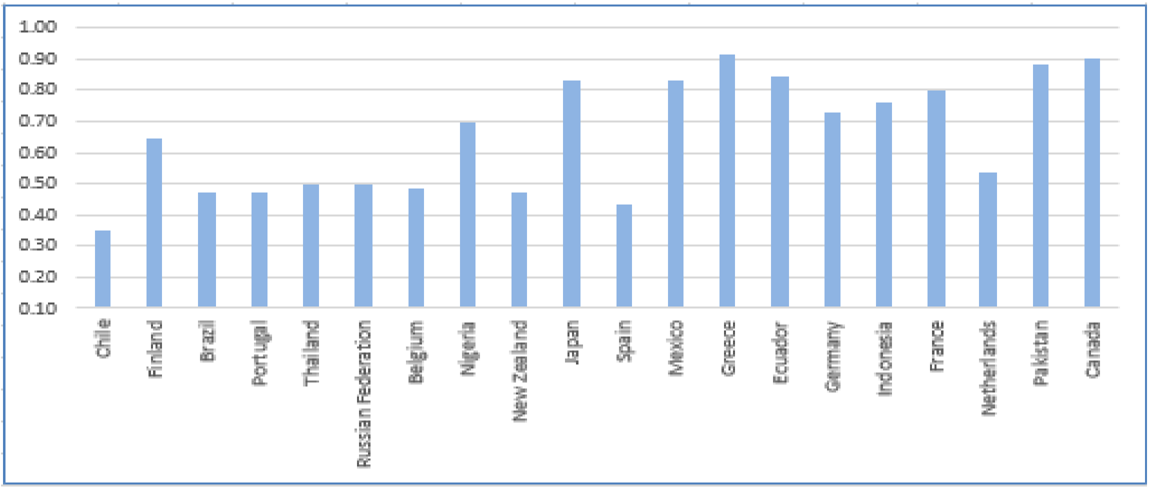

Figure 8. Standard deviation for friendliness of opinion leaders.

Results show that opinion leaders from Chile, Finland and Brazil had the most favourable views of Palestine. It is also noticed that countries that posted the most tweets; i.e., Canada, the US and the UK, are ranked low in terms of the friendliness of their opinion leaders. Looking at the standard deviations, the variation among opinion leaders increases when the friendliness score is low and *vice versa*. This indicates that opinion leaders were highly divided over the Palestinian-Israeli conflict. For example, the variance is high in countries like Germany and Canada in which the friendliness scores are low, while the variance is low in Chile and Spain.

Figure 9 shows the friendliness of opinion leaders as compared to the friendliness of the top twenty countries that generated the most tweets. In general, the attitude of opinion leaders looks consistent with the attitude of their countries for most countries. However, opinion leaders have a slightly more positive attitude towards Palestine as compared to the attitude of the public opinion as in the cases of the UK, Brazil and Chile. On the contrary, countries like Japan, France and China have leaders with less favourable views towards Palestine as compared to the country's friendliness score.

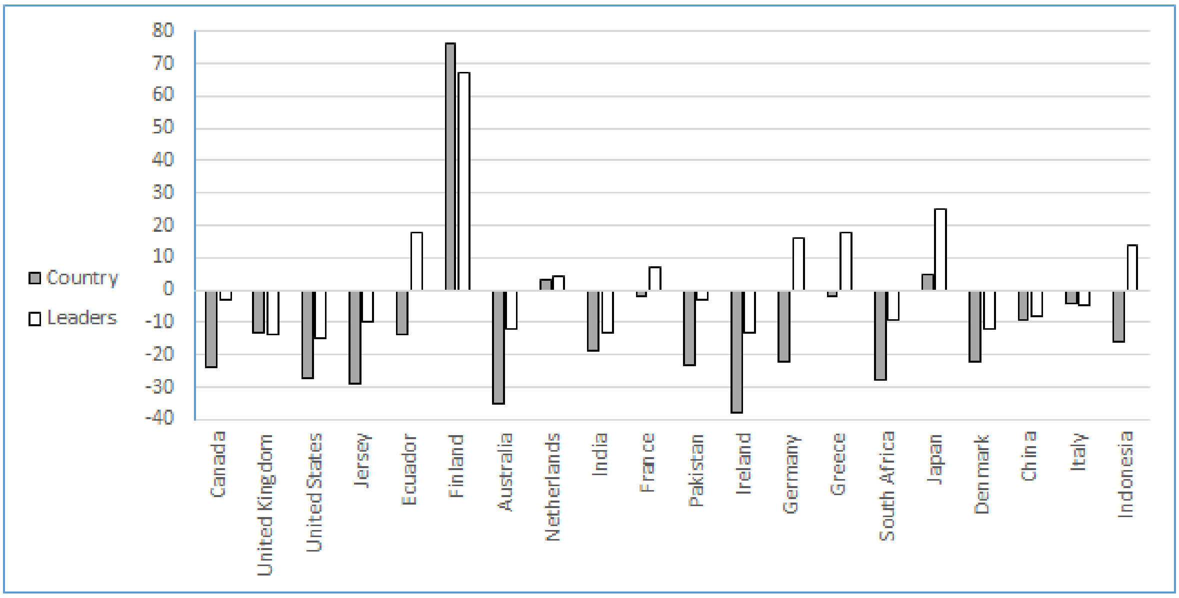

Figure 9. Friendliness of countries *vs.* opinion leaders.

9.2.2 Influence of Individual's Background

Individuals who share the same ethnicity, race or religion are likely to be sympathetic to each other's issues. For example, a large number of Muslim and Arab people living in Europe and North America provide continuous support to the Palestinian people. Part of this support comes through social networks in different forms, such as retweet campaigns, hashtags, fundraising and promoted tweets. The positive attitude of Muslim and Arab individuals is largely driven by shared culture or religious motivation.

The friendliness scores presented in Table 5 show the overall country's attitude, but do not show how this attitude is influenced by the background of Twitter users or how different groups, such as Muslims or Arabs, contribute to the public opinion in their countries. Identifying the attitudes of different groups will be helpful for decision makers and social media activists, so that they can alter their speech and dialogue according to the needs and motivates of each group.

For simplicity, we decided to classify Twitter users from each country into two groups based on their names: a group of people who have Arabic or Muslim names (we refer to it as "Arab_Muslims" group) and a group of people who have other names (we refer to it as "non-Arab_Muslim" group). One should note here that the group with Arabic names is not restricted to Arab people, but may include people from the wider Muslim world, such as Pakistanis, Iranians and Indians. Usernames can give an indicator of the ethnic or religious group to which a Twitter user belongs. However, the limitation of using usernames is that some Twitter users may use nicknames not related to their original names.

To identify Arabic names, we used a dictionary of Arabic names that we constructed from[viii]. In total, 828 Arabic names were included in the dictionary, besides the different ways of writing these names in both Arabic and English. Each Twitter username in the collected tweets was compared with the names in the dictionary. If the username contains an Arabic name, in either Arabic or English, it is added to the group of people with Arabic names. Otherwise, it is assumed to be a non-Arab user. After identifying users with Arabic and non-Arabic names, friendliness scores of each group is calculated.

Table 9 presents the analysis results for the top twenty countries that posted the most tweets. For each country, the friendliness score is presented along with the friendliness of the "Arab_Muslim" and "non_Arab_Muslim" groups. The percentage of tweets posted by each group is also shown. In general, the contribution of the Arab-Muslim communities was marginal for most countries as is evident from the small numbers of tweets posted by them. This result is expected, because Muslims and Arabs are minorities in most surveyed countries.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 05, No. 03, December 2019.

Users with Arabic names have more favourable views of Palestine. Friendliness was high among Arabs and Muslims in most major Western countries, such as Canada, the UK and the US. The friendliness scores were low or even negative in countries like Japan and South Africa. This result does not necessarily reflect the situation, because the number of tweets identified as being posted by Arabs and Muslims in these countries was too small to be representative of the entire population.

In general, the positive sentiment of Arabs and Muslims in most countries did not influence the overall public opinion due the small number of tweets. Apart from Finland, users with non-Arabic names have negative friendliness scores.

Table 9. Friendliness of user groups for the countries that posted the most tweets.

Country Name	Country's Friendliness	"Arab_Muslim" group		"Non_Arab_Muslim" group	
		Friendliness	Percentage of tweets	**Friendliness**	Percentage of tweets
Canada	-24.43	**50.76**	2.8%	**-26.65**	97.2%
United Kingdom	-13.31	**27.51**	1.7%	**-14.01**	98.3%
United States	-27.04	**34.51**	0.2%	**-28.64**	99.8%
Jersey	-29.16	**58.33**	0.1%	**-29.25**	99.9%
Ecuador	-14.88	**7.14**	0.45%	**-14.87**	99.55%
Finland	75.97	**20.00**	0.1%	**76.05**	99.9%
Australia	-35.46	**22.64**	1.7%	**-36.46**	98.3%
Netherlands	3.82	**7.14**	1%	**3.78**	99%
India	-19.38	**1.10**	6.3%	**-20.75**	93.7%
France	-1.40	**23.08**	2.1%	**-1.93**	97.9%
Pakistan	-23.89	**26.79**	21.52%	**-37.80**	78.48%
Ireland	-37.38	**12.50**	0.82%	**-37.80**	99.18%
Germany	-23.01	**3.33**	3.61%	**-24.00**	96.39%
Greece	-2.56	**9.52**	2.56%	**-2.88**	97.44%
South Africa	-27.75	**-2.70**	5.16%	**-29.12**	94.84%
Japan	5.61	**0.00**	0.31%	**5.63**	99.69%
Denmark	-22.69	**4.55**	3.44%	**-23.66**	96.56%
China	-9.12	**3.09**	15.25%	**-11.32**	84.75%
Italy	-4.85	**20.00**	2.6%	**-5.52**	97.4%
Indonesia	-16.31	**14.29**	2.7%	**-17.16**	97.3%

10. CONCLUSIONS AND FUTURE WORK

This research proposes an approach for political sentiment analysis at both country and individual levels. The approach was implemented to explore the international public opinion towards the Palestinian-Israeli conflict by using Twitter data. A dataset consisting of 178,524 tweets posted during 2016, was collected and pre-processed. The polarities of tweets were first measured by using a sentiment analyser that was specially trained to identify the sentiment about Palestine. Several features were then extracted and analyzed to provide a deep insight into the public opinion.

There are many directions to extend this work in the future: First, we aim to use a larger dataset of tweets that span over several years. This will likely generate more reliable and generalizable results. Second, we aim to improve the sentiment analyser by training it with a larger volume of tweets. This is crucial, because the whole analysis is based on the polarities generated by the sentiment analyser. Third, we aim to explore and use more reliable approaches to identify opinion leaders and individual's background and characteristics. Forth, we plan to perform content analysis by means of topic modelling in order to explore what people are discussing with respect to the Palestinian-Israeli conflict.

We think that other researchers, not necessarily from the IT discipline, can also build on these results to gain deeper insights. For example, the results from this analysis may be compared with the results of related national polls in order to explore similarities and/or differences.

REFERENCES

[1] A. Bermingham and A. Smeaton, "On Using Twitter to Monitor Political Sentiment and Predict Election Results," Proceedings of the Workshop on Sentiment Analysis Where AI Meets Psychology (SAAIP 2011), pp. 2-10, 2011.

[2] A. Tumasjan, T. O. Sprenger, P. G.Sandner and I. M. Welpe, "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment," Proceedings of the 4th International AAAI Conference on Weblogs and Social Media, Association for the Advancement of Artificial Intelligence, pp. 178-185, 2010.

[3] A. Jungherr, "Twitter Use in Election Campaigns: A Systematic Literature Review," Journal of Information Technology & Politics, vol. 13, no. 1, pp. 72-91, 2016.

[4] B. O'Connor, R. Balasubramanyan, B. R. Routledge and N. A. Smith, "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series," Proceedings of the 4th International AAAI Conference on Weblogs and Social Media, Association for the Advancement of Artificial Intelligence, pp. 122-129, 2010.

[5] H. Wang, D. Can, A. Kazemzadeh, F. Bar and S. Narayanan, "A System for Real-time Twitter Sentiment Analysis of 2012 US Presidential Election Cycle," Association for Computational Linguistics, pp. 115-120, 2012.

[6] E. Siapera, "Tweeting# Palestine: Twitter and the Mediation of Palestine," International Journal of Cultural Studies, vol. 17, no. 6, pp. 539-555, 2014.

[7] E. Siapera, G. Hunt and T. Lynn, "# GazaUnderAttack: Twitter, Palestine and Diffused War," Information, Communication & Society, vol. 18, no. 11, pp. 1297-1319, 2015.

[8] P. Sobkowicz, M. Kaschesky and G. Bouchard, "Opinion Mining in Social Media: Modeling, Simulating and Forecasting Political Opinions in the Web," Government Information Quarterly, vol. 29, no. 4, pp. 470-479, 2012.

[9] J. A. Balazs and J. D. Velásquez, "Opinion Mining and Information Fusion: A Survey," Information Fusion, vol. 27, pp. 95-110, 2016.

[10] E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Urena-López and A. R. Montejo-Ráez, "Sentiment Analysis in Twitter," Natural Language Engineering, vol. 20, no. 1, pp. 1-28, 2014.

[11] F. Bellini and N. Fiore, "Exploring Sentiment on Financial Market Through Social Media Stream Analysis," Reshaping Accounting and Management Control Systems, Springer, pp. 115-129, 2017.

[12] M. Nardo, M. Petracco-Giudici and M. Naltsidis, "Walking Down Wall Street with a Tablet: A Survey of Stock Market Predictions Using the Web," Journal of Economic Surveys, vol. 30, no. 2, pp. 356-369, 2016.

[13] J. Bollen, H. Mao and X. Zeng, "Twitter Mood Predicts the Stock Market," Journal of Computational Science, vol. 2, no. 1, pp. 1-8, 2011.

[14] P. N. Howard, G. Bolsover, B. Kollanyi, S. Bradshaw and L.-M. Neudert, "Junk News and Bots during the US Election: What Were Michigan Voters Sharing over Twitter," Computational Propaganda Research Project, Oxford Internet Institute, Data Memo, 2017.

[15] D. Murthy, "Twitter and Elections: Are Tweets, Predictive, Reactive or a Form of Buzz?," Information, Communication & Society, vol. 18, no. 7, pp. 816-831, 2015.

[16] P. L. Francia, "Free Media and Twitter in the 2016 Presidential Election: The Unconventional Campaign of Donald Trump," Social Science Computer Review, vol. 36, no. 4, pp. 440-455, 2018.

[17] D. Gayo-Avello, "No, You Cannot Predict Elections with Twitter," IEEE Internet Computing, vol. 16, no. 6, pp. 91-94, 2012.

[18] P. Howard, B. Kollanyi and S. C. Woolley, "Bots and Automation over Twitter during the Second US Presidential Debate," The Computational Propaganda Project, Oxford Internet Institute, pp. 1-4, 2016.

[19] D. Freelon and D. Karpf, "Of Big Birds and Bayonets: Hybrid Twitter Interactivity in the 2012 Presidential Debates," Information, Communication & Society, vol. 18, no. 4, pp. 390-406, 2015.

[20] K. Gorkovenko and N. Taylor, "Understanding How People Use Twitter during Election Debates," Proceedings of the ACM 31st British Computer Society Human Computer Interaction Conference, BCS Learning & Development, Ltd., pp. 88, 2017.

[21] D. Kreiss, "Seizing the Moment: The Presidential Campaigns' Use of Twitter during the 2012 Electoral Cycle," New Media & Society, vol. 18, no. 8, pp. 1473-1490, 2016.

[22] T. Lansdall-Welfare, F. Dzogang and N. Cristianini, "Change-point Analysis of the Public Mood in UK Twitter during the Brexit Referendum," Proc. of IEEE 16th International Conference on Data Mining Workshops (ICDMW), pp. 434-439, 2016.

[23] N. A. Diakopoulos and D. A. Shamma, "Characterizing Debate Performance *via* Aggregated Twitter Sentiment," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 1195-1198, 2010.

[24] S. Stieglitz and L. Dang-Xuan, "Political Communication and Influence through Microblogging: An Empirical Analysis of Sentiment in Twitter Messages and Retweet Behavior," Proc. of the 45th IEEE Hawaii International Conference on System Sciences, pp. 3500-3509, 2012.

[25] I. Alfina, D. Sigmawaty, F. Nurhidayati and A. N. Hidayanto, "Utilizing Hashtags for Sentiment Analysis of Tweets in the Political Domain," Proceedings of the 9th International Conference on Machine Learning and Computing, ACM, pp. 43-47, 2017.

[26] P. Grover, A. K. Kar, Y. K. Dwivedi and M. Janssen, "Polarization and Acculturation in US Election 2016 Outcomes–Can Twitter Analytics Predict Changes in Voting Preferences," Technological Forecasting and Social Change, pp. 1-23, 2018.

[27] R. Bose, R. K. Dey, S. Roy and D. Sarddar, "Analyzing Political Sentiment Using Twitter Data," Information and Communication Technology for Intelligent Systems, Springer, pp. 427-436, 2019.

[28] D. Paul, F. Li, M. K. Teja, X. Yu and R. Frost, "Compass: Spatio Temporal Sentiment Analysis of US Eelection What Twitter Says!," Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, pp. 1585-1594, 2017.

[29] O. Almatrafi, S. Parack and B. Chavan, "Application of Location-based Sentiment Analysis Using Twitter for Identifying Trends towards Indian General Elections 2014," Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication, ACM Edn., 2015.

[30] S. Vosoughi, H. Zhou and D. Roy, "Enhanced Twitter Sentiment Classification Using Contextual Information," arXiv preprint arXiv:1605.05195, 2016.

[31] J. Deegan, J. Hogan, S. Feeney and B. K. O'Rourke, "The Self and Other: Portraying Israeli and Palestinian Identities on Twitter," Irish Communication Review, vol. 16, no. 1, Article 8, 2018.

[32] J. W. Pennebaker, M. E. Francis and R. J. Booth, "Linguistic Inquiry and Word Count: LIWC 2001," Mahway: Lawrence Erlbaum Associates, vol. 71, 2001.

[33] P. Burnap, R. Gibson, L. Sloan, R. Southern and M. Williams, "140 Characters to Victory?: Using Twitter to Predict the UK 2015 General Election," Electoral Studies, vol. 41, pp. 230-233, 2016.

[34] A. Ceron, L. Curini and S. M. Iacus, "Using Sentiment Analysis to Monitor Electoral Campaigns: Method Matters—Evidence from the US and Italy," Soci. Sci. Comp. Review, vol. 33, no. 1, pp. 3-20, 2015.

[35] A. Ceron, L. Curini and S. M. Iacus, "To What Extent Sentiment Analysis of Twitter Is Able to Forecast Electoral Results? Evidence from France, Italy and the United States," Proc. of the 7th ECPR General Conference Sciences Po, Bordeaux, pp. 5-8, 2013.

[36] D. J. Hopkins and G. King, "A Method of Automated Non-parametric Content Analysis for Social Science," American Journal of Political Science, vol. 54, no. 1, pp. 229-247, 2010.

[37] H. T. Le, G. Boynton, Y. Mejova, Z. Shafiq and P. Srinivasan, "Revisiting the American Voter on Twitter," Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, pp. 4507-4519, 2017.

[38] F. Marozzo and A. Bessi, "Analyzing Polarization of Social Media Users and News Sites during Political Campaigns," Social Network Analysis and Mining, Springer, vol. 8, no. 1, 2018.

[39] S. Martin-Gutierrez, J. C. Losada and R. M. Benito, "Recurrent Patterns of User Behavior in Different Electoral Campaigns: A Twitter Analysis of the Spanish General Elections of 2015 and 2016," Complexity Jour., vol. 2018, 2018.

[40] E. M. Cody, A. J. Reagan, L. Mitchell, P. S. Dodds and C. M. Danforth, "Climate Change Sentiment on Twitter: An Unsolicited Public Opinion Poll," PloS One, vol. 10, no. 8, pp. e0136092, 2015.

[41] N. S. Khan, M. Ata and Q. Rajput, "Identification of Opinion Leaders in Social Network," Proc. of IEEE International Conference on Information and Communication Technologies (ICICT), pp. 1-6, 2015.

[42] F. Bodendorf and C. Kaiser, "Detecting Opinion Leaders and Trends in Online Social Networks," Proceedings of the 2nd ACM Workshop on Social Web Search and Mining, pp. 65-68, 2009.

[43] M. Solomon, R. Russell-Bennett and J. Previte, Consumer Behaviour, Pearson Higher Education AU, 3rd Edition, 2012.

[44] F. Riquelme and P. González-Cantergiani, "Measuring User Influence on Twitter: A Survey," Information Processing & Management, vol. 52, no. 5, pp. 949-975, 2016.

[45] M. Gaurav, A. Srivastava, A. Kumar and S. Miller, "Leveraging Candidate Popularity on Twitter to Predict Election Outcome," Proceedings of the 7th Workshop on Social Network Mining and Analysis, ACM, Article 7, 2013.

[46] DNOiSE., "Followthehashtag // Free twitter search analytics and business intelligence tool," [Online], Available: http://www.followthehashtag.com/.

[47] O. Owoputi, B. O'Connor, C. Dyer, K. Gimpel, N. Schneider and N. A. Smith, "Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters," Proceedings of NAACL-HLT, Association for Computational Linguistics, pp. 380–390, 2013.

[48] K. Gimpel, N. Schneider and B. O'Connor, "Annotation Guidelines for Twitter Part-of-Speech Tagging Version 0.3," [Online], Available: http://www.ark.cs.cmu.edu/TweetNLP/annot_guidelines.pdf.", 2013.

[49] M. J. Idzelis, "The Java Open Source Spell Checker," [Online], Available: http://jazzy.sourceforge.net/, [Accessed: 15-08-2019].

[50] Alias-i., "LingPipe 4.1.0.," [Online], Available: http://alias-i.com/lingpipe, [Accessed: 15-08-2019].

[51] C. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. Bethard and D. McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit," pp. 55-60, DOI:10.3115/v1/p14-5010, 2014.

[52] M. Thelwall, K. Buckley, G. Paltoglou, C. Cai and A. Kappas, "SentiStrength," [Online], Available: http://sentistrength.wlv.ac.uk/, [Accessed: 15-08-2019].

[53] R. Kumar, H. S. Pannu and A. K. Malhi, "Aspect-based Sentiment Analysis Using Deep Networks and Stochastic Optimization," Neural Computing and Applications, pp. 1-15, 2019.

[54] A. Valdivia, M. V. Luzón, E. Cambria and F. Herrera, "Consensus Vote Models for Detecting and Filtering Neutrality in Sentiment Analysis," Information Fusion, vol. 44, pp. 126-135, 2018.

[55] A. Milani, N. Rajdeep, N. Mangal, R. K. Mudgal and V. Franzoni, "Sentiment Extraction and Classification for the Analysis of Users' Interest in Tweets," International Journal of Web Information Systems, vol. 14, no. 1, pp. 29-40, 2018.

[56] E. Kušen and M. Strembeck, "Politics, Sentiments and Misinformation: An Analysis of the Twitter Discussion on the 2016 Austrian Presidential Elections," Online Social Networks and Media, vol. 5, pp. 37-50, 2018.

[57] M. Bouazizi and T. Ohtsuki, "Sentiment Analysis: From Binary to Multi-class Classification: A Pattern-based Approach for Multi-class Sentiment Analysis in Twitter," IEEE Access, vol. 5, pp. 1-6, 2016.

[58] U. Yaqub, S. Chun, V. Atluri and J. Vaidya, "Sentiment-based Analysis of Tweets during the US Presidential Elections," Proceedings of the 18th Annual International Conference on Digital Government Research, ACM, pp. 1-10, 2017.

[59] Apache.org, "Apache Spark: Unified Analytics Engine for Big Data," [Online], Available: https://spark.apache.org/, [Accessed: 15/08/2019].

[60] L. Bornmann and R. Haunschild, "How to Normalize Twitter Counts? A First Attempt Based on Journals in the Twitter Index," Scientometrics, vol. 107, no. 3, pp. 1405-1422, 2016.

[61] R. Taylor, "Interpretation of the Correlation Coefficient: A Basic Review," Journal of Diagnostic Medical Sonography, vol. 6, no. 1, pp. 35-39, 1990.

[62] E. Dubois and D. Gaffney, "The Multiple Facets of Influence: Identifying Political Influentials and Opinion Leaders on Twitter," American Behavioral Scientist, vol. 58, no. 10, pp. 1260-1277, 2014.

[63] W. W. Xu, Y. Sang, S. Blasiola and H. W. Park, "Predicting Opinion Leaders in Twitter Activism Networks: The Case of the Wisconsin Recall Election," American Behavioral Scientist, vol. 58, no. 10, pp. 1278-1293, 2014.

[64] D. S. Moore, G. P. McCabe and B. A. Craig, Introduction to the Practice of Statistics, 7th Edition, W.H. Freeman & Company, 2012.

ملخص البحث:

في ظل النمو الهائل لمنصات التواصل الاجتماعي، لا ينشر الناس معلومات عامة فحسب، وإنما ينشرون آراءهم السياسية أيضاً. وقد استخدم الكثير من الأبحاث محتوى وسائل التواصل الاجتماعي لتحليل الرأي العام تجاه الأحداث السياسية والتنبؤ به. يقدم هذا العمل دراسة تحليلية لقياس الرأي العام السياسي تجاه النزاع الفلسطيني – الإسرائيلي باستخدام بيانات تويتر. تستخدم هذه الدراسة نموذجاً مبتكراً لتحليل البيانات يُعنى بمستويين من التحليل هما: التحليل على مستوى الدولة، والتحليل على مستوى الفرد. يهدف التحليل على مستوى الدولة الى استكشاف الاتجاه الإجمالي للدّول تجاه فلسطين، وذلك عبر: 1. تحديد الدول التي صدر منها أكثر التغريدات المتعلقة بالموضوع؛ 2. قياس دراجة الصداقة لكل دولة نحو فلسطين؛ 3. تحليل التغير في الرأي مع الوقت. أما التحليل على المستوى الفردي فيهدف الى تحليل البيانات بناءً على نشاط الأفراد وخلفياتهم. وقد جرى تحليل اتجاهات كل من قادة الرأي والمجموعات الإثنيّة ومناقشتها في ضوء اتجاهات الدول.

إن التجربة الغنية التي يقدمها هذا البحث من خلال النموذج المقترح، والإجراءات التي اتُّبعت خطوة بخطوة، وتنوع تقنيات التحليل، ومناقشة النتائج، من شأنها أن تضع معلومات وافرة بين أيدي مطوري النماذج ومحللي البيانات المهتمين في تحليل الآراء التي يعبر عنها على منصات التواصل، فيما يتعلق بالنزاعات السياسية على وجه الخصوص.

(i) Google Trends, https://trends.google.com/trends/?hl=en

(ii) http://www.aljazeera.com/news/2016/07/israeli-demolitions-displace-dozens-palestinians 160713124336539.html

(iii) http://www.independent.co.uk/news/world/middle-east/israel-border-crossing-checkpoint-palestinians-west-bank-btselem-a7106486.html

(iv) http://www.btselem.org/press_releases/20160727_house_demolitions_in_area_c

(v) http://www.independent.co.uk/news/world/middle-east/jerusalem-mayor-palestinians-animals-terror-attack-two-killed-meir-turgeman-a7355116.html

(vi) https://www.middleeastmonitor.com/20161013-unesco-vote-no-link-between-al-aqsa-and-judaism/

(vii) http://www.aljazeera.com/news/2016/10/palestinians-unesco-vote-al-aqsa-compound-161015133808135.html
(viii) Behind the Name: the etymology and history of first names, https://www.behindthename.com (Last Access: 15/04/2019)