Ahmad Salameh

SPREADING IDEOLOGIES THROUGH TWEETS: EXAMINING EXTREME AND MODERATE MUSLIMS USAGE OF TWITTER

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION SYSTEMS
2018

ABSTRACT

Salameh, Ahmad
Spreading Ideologies through Tweets: Examining Extreme and Moderate Muslims Usage of Twitter
Jyväskylä: University of Jyväskylä, 2018.
Information Systems, Master's Thesis
Supervisors: Zhang, Yixin (Sarah) and Semenov, Alexander

Twitter enables groups with certain agendas to organize and distribute their ideologies. This research compares the different practices performed by the extreme and the moderate Muslims to build their networks and recruit more followers. We carry out our research in the context of religious communication on Twitter. The study contains two data sets of tweets written in the Arabic language; the first one is retrieved from certain accounts that are leaning to extremism or moderation based on the generated content by the account holders, and the second one is obtained through predefined keywords related to Islam,. Collected tweets and retweets were analyzed through network analysis to understand users' networking behavior, and to examine whether polarization exists. Regression analysis showed that negative sentiment in tweets has a significant positive impact on the retweeting quantity, while interestingly; positive sentiment was not statically significant to affect retweeting. Features of hashtags, URLs, tweet length, and the number of followers have a positive effect on retweeting, whereas the number of mentioned names has a significant negative effect on retweeting. Through network and centrality measures, we found that extreme and moderate users have higher frequency of interaction within their ideological group than with the ideologically-opposed users. We also suggest, based on our findings and related research, that extreme and moderate Muslims use provocation to introduce their partisan content to ideologically-opposed users.

Keywords: Twitter, Arabic Tweets, Extreme, Moderate, Muslims, Information Diffusion, Network Analysis.

FIGURES

TABLES

TABLE OF CONTENTS

1 INTRODUCTION

Today, with the help of easy access to internet, more and more people have shifted from connecting with other people personally towards building networks that allow communication through social media platforms such as Facebook, Twitter, YouTube, etc. For example, Twitter's average monthly active users, have grown to 328 million in the first quarter of 2017 adding more 9 million users to the previous quarter (Forbes, 2017). Stieglitz and Dang-Xuan (2013) stated that this enormous expansion in user base in social media has affected communication and debate in our modern society. This could be clearly seen in the cross-ideological discourse between different groups on such social-media platforms. This research will compare the usage of Twitter by two ideologically-opposed groups; extreme Muslims and moderate ones.

1.1 Why Social Media?

In the past, communicating with large amount of people was much more difficult than today; it was only done by few people who had enough power of technical infrastructure to reach people. Nowadays, social media made it much easier to connect with other people from all around the world, and it has become widely used among different populations and cultures. The number of social media users in 2017 is 2.46 billion users (Statista, 2017), which is more than third of the world population; 33% out of 7.5 billion (UN, 2017). Statista (2017) added that this number is estimated to grow in 2019 to 2.77 billion and is expected to be ever-increasing worldwide.

Social media has expanded into being the prevalent type of media to share and distribute information. Stieglitz and Dang-Xuan (2013) argued that "mainstream adoption of social media" and "widespread access to the Internet" has made social networks and weblogs most dominant in information dissemination. They also discussed how short-content microblogging (i.e. phrases, quick comments, images, or links to videos) has been adopted by online users to share news, advocate for political stands, practice marketing, and follow real-time events.

1.2 Why Muslims, from Arabic-speaking Countries?

Statistics about religious population of the world are showing that Islam is now the fastest growing religion in the world (Huda, 2017). The recent years have witnessed significant concern about what is and what is not Islam; in their book "Framing Muslims", (Morey and Yaqin, 2011) described how the headlines on front pages and television screens scream out at us every day to draw a certain picture about Muslims. Extreme Muslims see an opportunity to apply their radical values in extreme groups and present it to the media while approving and adopting many terrorist attacks, spreading fear and panic through all over the world. While, on the other hand, moderate Muslims try to present their case to the masses using different channels of media as well; that this is not the way Islam should be, instead, the word of God (Allah) should be spread by love and peace, leading many people to believe that there is a different face of Islam that should be considered, rather than just fight Islam and Muslims in general.

Studies about Muslim's usage of social media are usually made with generalizations of all Muslims regardless of their type and where they come from. This research is going to focus on Muslims from the Arabic-speaking countries "from the Middle East and North Africa". The number of Muslims in the Arab world is more than 370 million which make up nearly a quarter "21%" of the total Muslim population of the world "1.8 billion" (Huda, 2017). Muslims of the Arab world speak Arabic as their native language which is the language used in the Quran; "the central religious text of Islam" (Jones, 2011), and Sunnah; the Islamic teachings and practices taught by the Islamic prophet Muhammad (Islahi, 2011). In her article about the importance of the Arabic language in Islam, Huda (2017) stated that "it is the Arabic language that serves as the common link joining this diverse community of believers and is the unifying element that ensures believers share the same ideas".

1.3 Why Twitter?

Alongside several other microblogging services, Twitter stands as one of the major social networking services by being ranked the fourth after Facebook, YouTube, and Instagram respectively, with 330 million monthly active users (Kallas, 2017). Conover et al. (2011) stated that Twitter is a prevalent social networking and microblogging website, in which, users have the ability to post 140-character short messages "tweets", allowing them to find and share matters of interest in a network that is built with "Followers" in real time. Stieglitz and Dang-Xuan (2013) defined "following" on Twitter as subscribing to a user's tweets and described this action as being not automatically mutual; with the user being followed is having the option to follow back or not.

Users on Twitter can interact with each other by "user-accepted norms", they can answer some questions about an issue or draw attention to an external matter by starting the conversation through using the "@" sign in order to indicate the receiver of the message (Stieglitz & Dang-Xuan, 2013). In addition, Twitter users can communicate via two main public methods; retweets and mentions. Conover et al. (2011) defined retweeting as "a form of endorsement, allowing individuals to rebroadcast content generated by other users, thereby raising the content's visibility" (as cited in Boyd, Golder, and Lotan, 2008). On the other hand, mentions operate in a different manner; they enable a user to write the name of another user directly in the public feed or to point out an individual in the third person (Conover et al., 2011 as cited in Honey and Herring 2008).

FIGURE 1 Tweet with retweet and mention features

1.4 Twitter and Social Change

Twitter and other social media platforms (i.e. Facebook, YouTube) play a major role alongside in social change. Oh et al. (2015) concluded that Twitter had a crucial part in social change that was presented in the 2011 Egypt Revolution and forced President Mubark to resign from his 30-year dictatorship. Twitter enables groups with certain agendas to organize and distribute their ideologies through tweets. Conover et al. (2011) investigated how Twitter facilitated the communication process in the network between "ideologically-opposed individuals" with different political orientations. Similarly, this study will follow Muslims' usage of the social network, and compare between both types; extreme Muslims (i.e. ISIS, Boko Haram, the Taliban and Al-Qaeda - all of these groups follow Wahhabism, an extreme conservative branch of Islam - according to the Global Terrorism Index, 2016), and the moderate ones (i.e. Liberal and Progressive Muslim Movements, according to Safi, 2003) to build their networks and recruit more followers. Therefore, the aim of the study is to answer the research question:

RQ: *how extreme Muslims and moderate ones are connected on Twitter and how do they disseminate their ideologies on the platform?*

2 RELATED LITERATURE AND HYPOTHESES DEVELOPMENT

2.1 Users Network and Information Diffusion

In the research paper “Political Polarization on Twitter”, Conover et al. (2011) analyzed 250,000 tweets to examine how political Communication Networks are formed on Twitter. They defined how many nodes are there in their “largest connected component accounts”; and then only focused on these accounts for the rest of the analysis because of their dominance in the network. They performed their analysis through different stages; they used clustering algorithms of the network to explore the two different communities, statistical analysis of tweets’ content to present that generated tweets by users of the same community have more similar content than those generated by users from different communities.

To examine the community structure for our network, we used the methodology applied in the study of Ji et al., (2015), in which they established articles relationship and visualized the article network using an analytical and graphical tool, Gephi; which was described as an open source software for graph and network analysis. It helps data analysts to understand graphs and reveal hidden patterns and test their hypotheses (Gephi.org, 2017). Ji et al., (2015) analyzed networks using graph diameter, closeness centrality, and module classes to test the relationships between different articles. They revealed the distribution of articles and how they formed aggregations in specific module classes. Using similar evaluation measures, we will use the network graph measures to identify extreme and moderate communities and provide mathematical support for our hypothesis.

As Conover et al. (2011) used manual annotating of users to determine which ideology the belong to, in our study, we conducted a questionnaire to label users as extreme, moderate, or neutral, based on their generated content (discussed in detail in the data collection section). By focusing on extreme and moderate users, we could investigate their communities by using network centrality measures as in Ji et al., (2015), to test how they are connected within their community and with the ideologically-opposed one.

Based on the findings of Conover et al. (2011), we hypothesize that our sample users interact similarly with both; their own community and the opposed one. Conover et al. (2011) results showed that users retweet other users with similar opinions, but mention users with opposed ideologies. As retweeting is the key technique for information diffusion in Twitter, and mentions have negative association with retweeting (Suh et al., 2010), we suggest that users diffuse their information in a greater rate within their community rather than while interacting with external community users who believe in a different ideology. This is demonstrated by the following hypothesis:

H1: *Users' interaction with ideologically-similar users has a higher frequency of retweets and mentions than their interaction with the ideologically-opposed users.*

2.2 Tweet & User Features and Information Dissemination

Information diffusion in social networks has attracted many researchers to investigate its powerful capacity to direct or influence behavior of others or course of events. Many theories about information diffusion in social networks have been established by information systems researchers in different fields; for example physical and computational sciences, and for different reasons; such as specifying political-communication behavior on social networks, designing advertisements for social media users (Conover et al. 2011, and Stieglitz & Dang-Xuan, 2013), defining the crucial role of social ties for information dissemination that leads to forming opinions and discovering products (Susarla et al., 2012), and documenting the relationship between social interactions and the levels of similarities of users' generated content (Zeng, & Wei 2013).

A significant number of studies have concentrated on Twitter because the information diffusion is clearly represented on the platform through retweets, and it shows the different links in the social network by specifying which of them play major and minor roles in the information dissemination process (Stieglitz & Dang-Xuan, 2013). Past literature discussed the quantity and speed of retweeting (Yang & Counts 2010) and their relationship with virality and susceptibility in Information diffusion (Hoang & Lim2012). Social ties and users' status and their effects on information diffusion were also taken into consideration in previous research (Zeng, & Wei 2013) as well as the structural position in the social network (Susarla et al., 2012).

2.2.1 Emotions

Stieglitz and Dang-Xuan (2013) argued that literature about users' cognitive and arousal-related effects from written communication on sharing behavior could also work in the computer-mediated communication (CMC) presented in social media and particularly for their study on Twitter. They collected tweets from Twitter during the period of German elections due to the higher level of users' participation in political discourse on Twitter. The tweets that were collected contained the six most important German parties' names, after that, redundant and irrelevant tweets were eliminated. They conducted Sentiment analysis because it is a "systematic computer-based analysis of written text… It provides a fine-grained examination that aims to establish the overall orientation (positive or negative) and intensity (weak or strong) of the sentiments expressed by statements".

"SentiStrength" is the tool that was used to perform the sentiment analysis because of its ability to classify emotions in short informal messages from Myspace and Twitter"; it classifies positive and negative emotions with defined scales. Similarly, in this research, tweets are going through a Python code that contains a large lexicon of words, phrases, and emojis with already defined polarity (positive +1, negative -1). The code works in a similar way to SentiStrength; by granting the tweet a total polarity result based on the used words in it. Some examples are provided below to show how the code works with defining polarity of tweets.

FIGURE 2 Example for a tweet with positive-sentiment polarity

Tweet Translation in Figure 2: you know brother Jihad thank God and thanks for your book One Hundred Great Muslims, I am now interested in reading and in the future I can start writing. The code generated a positive-sentiment polarity, +5.

FIGURE 3 Example for a tweet with negative-sentiment polarity

Tweet Translation in Figure 3: for sure it is not freedom for a woman to go out in Uber without a mahram (legal male-escort of a woman) and this is haram (forbidden) but people who do it do not say it is. The code generated a negative-sentiment polarity, -7.

In addition, as presented in Stieglitz and Dang-Xuan (2013) and Zhang and Zhang (2016) studies, regression analysis will be performed to test the positive relationship between tweets' sentiments and emotions, and the retweet quantity. The main finding in Stieglitz and Dang-Xuan (2013) study was that "emotionally charged Twitter messages tend to be retweeted more often and more quickly compared to neutral ones". Thus, we hypothesize:

H2a: *Tweets that are emotionally-charged with positive sentiments have a higher likelihood to be retweeted comparing with tweets which are emotionally neutral.*

H2b: *Tweets that are emotionally-charged with negative sentiments have a higher likelihood to be retweeted comparing with tweets which are emotionally neutral.*

2.2.2 Tweet-related Features

Tweets' Textual Analysis will be conducted to find out how the tweets' textual content affects the retweeting behavior. Suh et al. (2010), Stieglitz and Dang-Xuan (2013), and Zhang and Zhang (2016) found out that there are certain tweet-related factors that affect the retweeting attitude; content features such as hashtags and URLs have a solid relationship with other tweet characteristics could also affect retweeting such as tweet length (number of characters), and number of mentioned names using the @ symbol.

FIGURE 4 Tweet with URL and mention features

Also, it was noticed that in which type of language the tweet written (i.e. formal or spoken language) is important. We believe that Modern Standard Arabic (MSA) affect the retweeting behavior because it is a closer type of language to Classical Arabic which is the language in the old Islamic texts and it is the literary standard language across the Arab world. However, there is no tool to detect MSA so far, but it could be manually done only on a small set of tweets (i.e. top 100 retweeted tweets) to figure out if the usage of MSA in tweet has a positive effect on retweeting quantity.

Based on the above discussed features, we hypothesize the following:

H3a: *Hashtags, have a positive effect on retweeting.*
H3b: *URLs have a positive effect on retweeting.*
H3c: *Tweet length has a positive effect on retweeting.*
H3d: *Number of mentioned names has a positive effect on retweeting.*

2.2.3 User-related Features

Susarla et al. (2012) examined the dynamics of diffusion of digital content constructed by a social network, and tested if the initial phase of the diffusion process is affected by different factors from those in the phases that follow. Past literature concluded that product diffusion is classified between aware-early adopters and late adopters. The authors concluded that the effect of central channels comes from their structural position in the social network, and the diffusion is processed through direct links in the network in this type of community. Features such as the number of followers and followees also affect the retweeting behavior. Stieglitz and Dang-Xuan (2013) added that it is likely that a user's followers have similar interests so it is expected that they will retweet their content. This could be even more powerful when users are following a religious ideology such as extreme or moderate Islamic ideology. This leads us to form the following hypothesis:

H4*: Number of followers has a positive effect on retweeting.*

3 METHODOLOGY

3.1 Data Collection

Two datasets were collected to test the hypotheses. From the first dataset, data was retrieved to test H1. Out of the second dataset, data was retrieved to test H2, H3, and H4.

3.1.1 The First Dataset

The first data set is collected from certain Twitter accounts that were classified as leaning to extremism or moderation based on their generated content. The labeled users are important to perform the network analysis as they will represent their community's ideology, and then we will be able to perform measurements on different communities to show how users are connected in the network (i.e. how extreme Muslims are connected within their community and with another community that has moderate Muslims).

To identify which accounts belong to extreme groups and which accounts belong to moderate ones, a classification process started by selecting random tweets (a tweet per account, 46 in total from 46 accounts without mentioning the account user or what they are leaning to). The tweets were then put into a questionnaire to be examined by a group of people (22 persons) with Islamic backgrounds to vote whether the tweet is extreme, moderate, or neutral in order to specify whether the holder of the account is leaning to extremism or moderation. If the voting was dominant (has more than 55% votes) for a certain category (Extreme, Moderate, Neutral), it is considered descriptive for an account, if not, the account will be discarded.

Each selected account has a number of followers between 1000 and 50,000. This will ensure that:

- The account holder has a sufficient number of followers that can be tracked when they interact with them.
- The account holder is not a very popular or famous person so followers will not interact with them only because of their status.

The questionnaire had 46 tweets; for each tweet there were three descriptions with a possibility to choose only one of them as follows:

- Extreme: the tweet and the account holder lean to extremism.
- Moderate: the tweet and the account holder lean to moderation.
- Neutral: If the person is not sure about their decision.

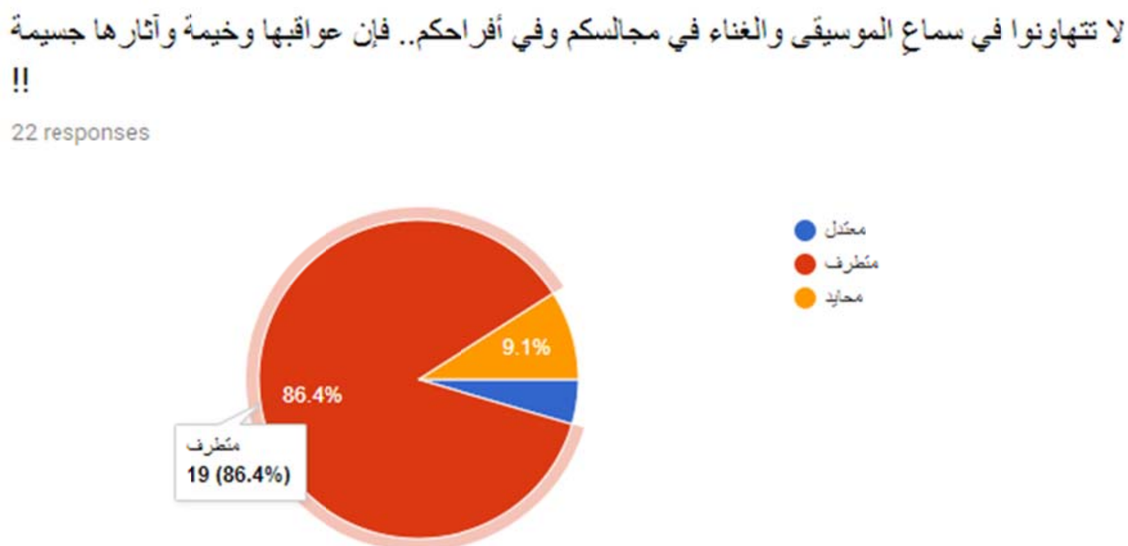

FIGURE 5 Example from the questionaire (extreme option as dominant)

Tweet translation in Figure 5: do not neglect hearing music in any environment.. Its consequences are dire and its effects are gross!!

Votes: 86.4% extreme (19 votes), 4.5% moderate (1 vote), 9.1% neutral (2 votes). Most of the sample voted for the tweet as extreme, thus, the extreme option is dominant. This indicates that the account holder leans to extremism.

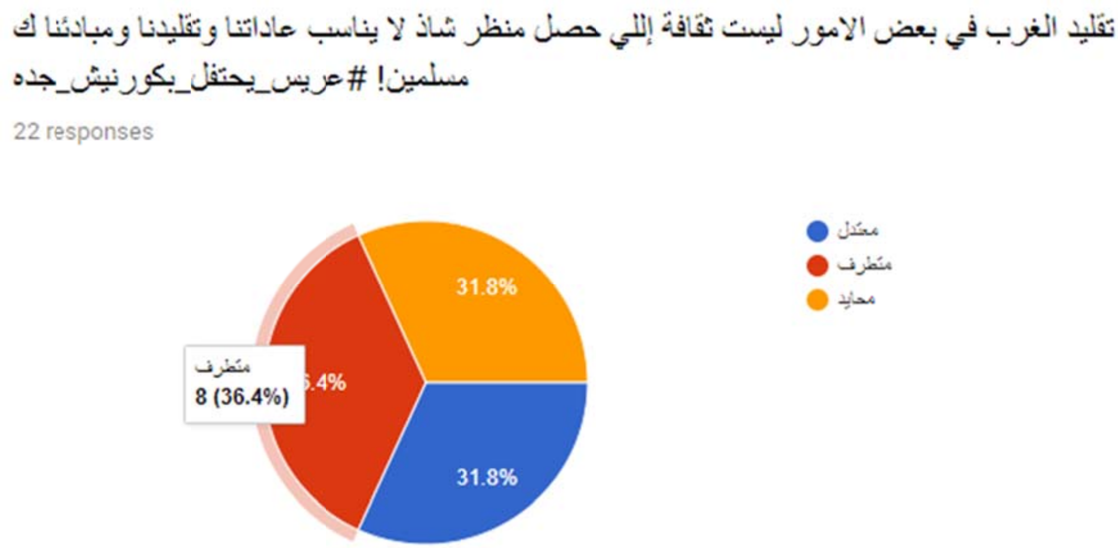

FIGURE 6 Example from the questionnaire (no dominant option)

Tweet translation in Figure 6: following the West in some things is not a culture, what happened is abnormal and does not fit our customs and traditions and principles as Muslims! #_groom_celebrates_jeddah

Votes: 36.4% extreme (8 votes), 31.8% moderate (7 votes), 31.8% neutral (7 votes). The voting is nearly divided between the three options with no dominant one. This cannot indicate whether the account leans to extremism or moderation, thus no data will be collected from such an account.

Based on the selection process above, 33 accounts were selected for the study. 16 were leaning to extremism and another 17 to moderation. The questionnaire results can be found in the Appendix 1.

3.1.2 The Second Dataset

The second dataset was retrieved based on specific keywords that are related to Islam (listed in Table 1 below). Collecting tweets based on keywords is one of the common data collection methods in social-media relevant research, for example, the method is adopted in Oh, Agwarl and Rao (2013) when examining rumor disseminations in social crisis, and in Conover et al. (2011) when examining political polarization on Twitter. The selected keywords referred to hot and sensitive topics that Arab Muslims are concerned and tweet about.

The program that was used to retrieve tweets is Python; collecting data through Python was done using Twitter Streaming API (Application Programming Interface) to download tweets. Four API keys were needed from Twitter to access the Twitter Streaming API; API key, API secret, Access token and Access token secret. Next, a Python library called "Tweepy" was used to connect to Twitter Streaming API and download the data. After that, a file was created to retrieve the data, it included the credentials from Twitter, specific keywords that are related to Islam, and the information we wanted to retrieve for our research; Retweets, User ID, Followers' Count, Friends' Count, Posts' Count, Location, Time Stamp, Hashtags, and Media URLs.

The program was run for two weeks to get a significant sample size. The data was stored in a text file in JSON format which makes the data easier to understand for humans, and then it was imported into Excel to present the data in columns and rows to simplify data measurement and analysis. 100,000 tweets were collected using this process over the two-week period. As not every tweet has been retweeted, the second dataset size was reduced from 100,000 tweets to 53,271 tweets, each of which has generated at least one retweet.

TABLE 1 Keywords used in Python to retrieve Twitter data

Keywords in Arabic	**English translation**
قيادة المرأة	Women to drive
حقوق المرأة	Women rights
نقاب	Niqab (garment of clothing that covers the face)
سعوديات نطلب	Saudi women
عمل المرأة	Women's right to work
سينيما السعودية	Cinemas in Saudi Arabia
يهود	Jews
دواعش	People who support Isis
الليبرالية	Liberalism
جهاد	Jihad, refers to armed struggle against unbelievers
حرام	haraam (taboo, forbidden)
وهابية	Wahhabism, fundamentalist Islamic movement
شيعة	Shia, a branch of Islam
روافض	Rafida, the term is used in a derogatory manner by Sunni Muslims who refer to Shias

نواصب	Nasibi, the term is used in a derogatory manner by Twelver Shias against Sunnis
طائفية	Sectarianism, form of bigotry, discrimination
رجم	Stoning
إرهاب	Terrorism
ختان	Circumcision
أمريكا	America

3.2 Measurements

For the data-analysis part of the study, network evaluation measures and multiple variables were determined to find out how Muslims interact within their group and with outsiders, and to measure the information diffusion through determining retweet quantity due to sentimental speech, tweet-related features, and user-related features.

3.2.1 First Dataset Measures

3.2.1.1 Network Evaluation Measures

From the second data-collection phase, two lists of accounts were selected to represent extreme Muslims (16 users) and moderate ones (17 users). The selected users were then imported to Gephi, which has a Twitter plugin (Twitter Streaming Importer) as shown in Figure 7 below. Levallois and Totet (2017) stated that the plugin allows the user to collect tweets in real time based on the topic that was chosen, acquire the mentioned users in these tweets and the connections between them, and visualize these connections after in Gephi or export the data to Excel. They added that the plugin has three main ways to collect tweets and user connections; the first way is by using the *Words to follow* tab which enables following one or multiple words, second is the *Locations to follow* tab that enables following the activity of one or multiple locations so any geo-tagged tweet will be captured. For our study, we chose the *Users to follow* tab which enables following the activity of one or multiple users including tweeting, retweeting, and mentioning the user. Twitter users that were collected in the first dataset were added into the *Users to follow* tab to capture their activity and connectivity.

After importing users to into the *Users to follow* tab, we had to select which Network Logic would be practical for our research. Levallois and Totet (2017) explained the network logic as how the incoming tweets are treated and how they are transformed into a set of nodes and edges. The dropdown list has 4 network logics to choose from; the first is *Full Twitter Network* which represents all types of entities (User, Tweet, Hastags, URL, etc...) as a graph, the second is *Hashtag Network* which creates hashtags network, the third is *Emoji Network* which functions in a similar way as the hashtag network but with focus on Emoji characters, and the fourth option is *User Network* which was suitable for our study because it captures the interaction between users represented by re-tweets and mentions, and the size of the edge which demonstrates the number of interactions between users.#

Next, we were connected to Twitter network, Gephi started collecting data and made nodes and edges based on users transactions in the network. After a 4-day collection period, 1,575 nodes and 2,713edges were created, which was enough to represent already-labeled users (extreme or moderate) networking behavior.

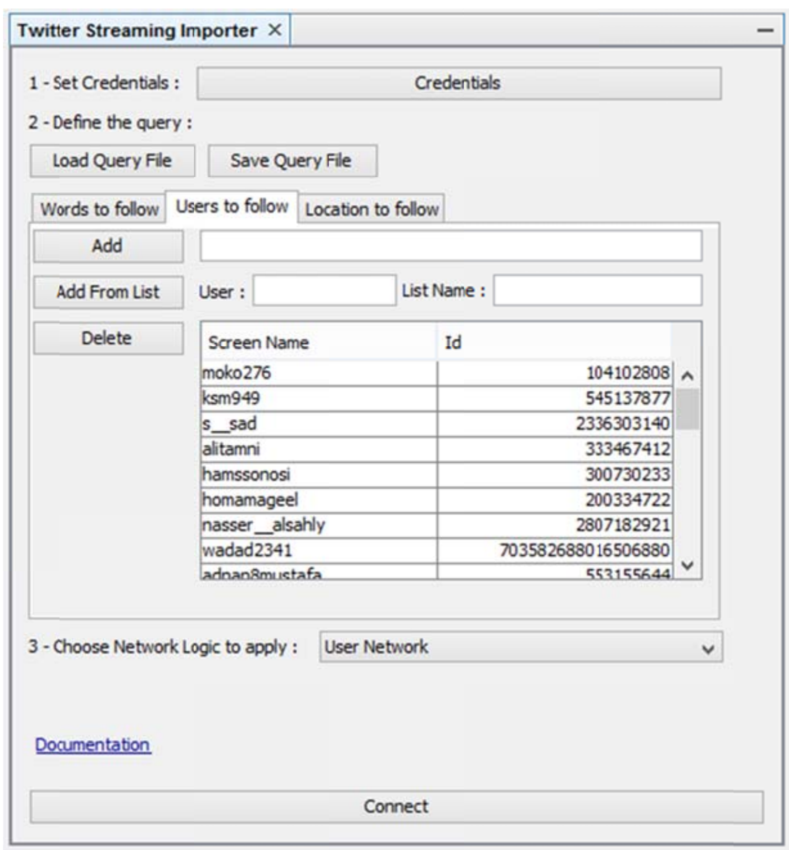

FIGURE 7 Twitter Streaming Importer plugin in Gephi

Network Centrality Measures

In order to mathematically represent the clustering of users, network diameter and betweenness centrality were measured. Network diameter is the shortest path between the most distant nodes. The diameter is representative of the linear size of a network. Running the network diameter in Gephi generates the following measures, as seen in Figure 8 below and defined by Hirst (2010):

Betweenness centrality: is measured by the number of shortest paths between any two nodes that pass through a particular node. Nodes around the edge of the network would typically have a low betweenness centrality. A high betweenness centrality might suggest that the individual is connecting various different parts of the network together.

Closeness centrality: shows how close a node is to all the other nodes in a network. A high closeness centrality means that there is a large average distance to other nodes in the network, and a small closeness centrality means there is a short average distance to all other nodes in the network.

Eccentricity: measures the distance between a node and the node that is furthest from it; so a high eccentricity means that the furthest away node in the network is a long way away, and a low eccentricity means that the furthest away node is actually quite close.

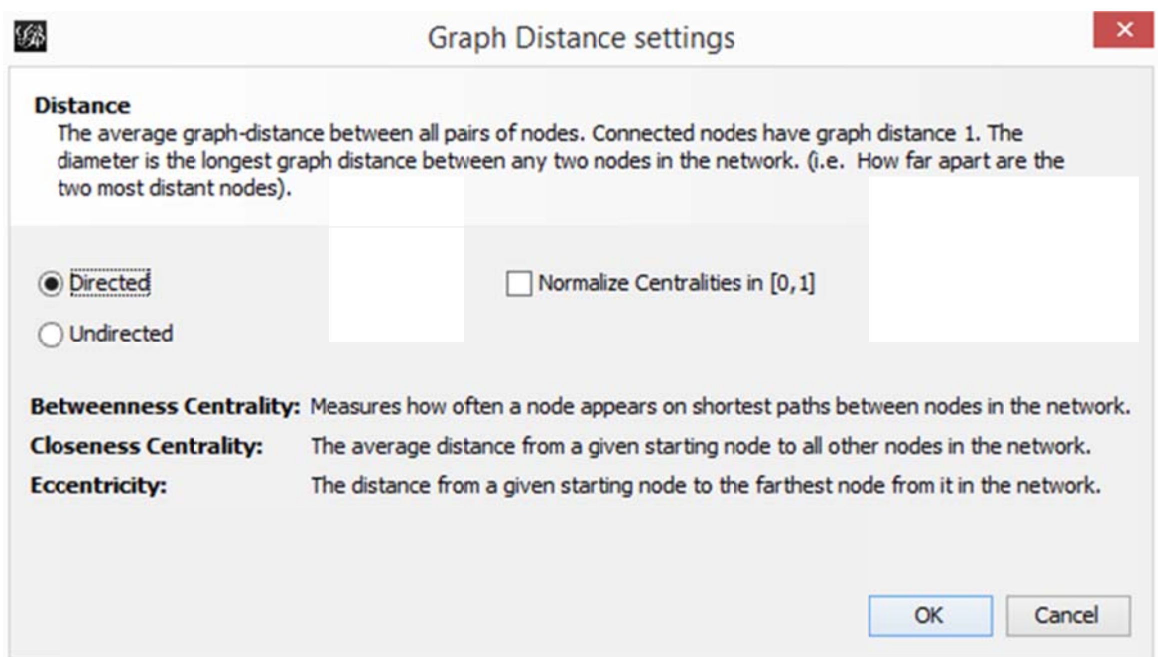

FIGURE 8 Measuring network diameter in Gephi

Modularity class

Ji et al., (2015) defined modularity class as a measure to detect communities in networks and determine the strength of divisions inside a network. They used this measure to investigate the resulting communities or module classes. This could be also performed by Gephi as shown in Figure 9 below.

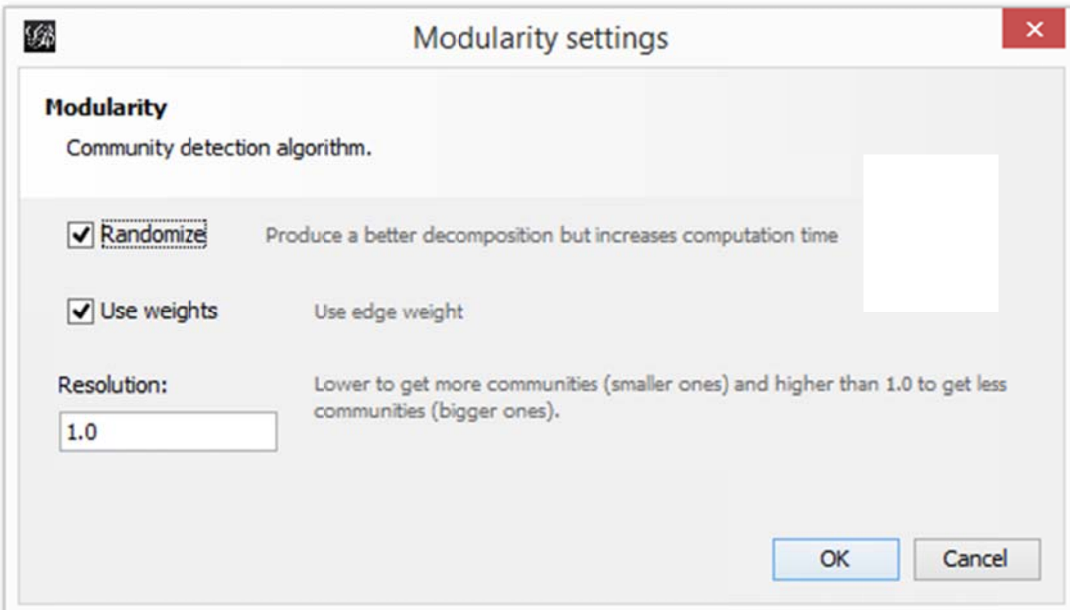

FIGURE 9 Measuring network modularity in Gephi

3.2.2 Second Dataset Measures

3.2.2.1 Dependent Variables

- *RtwtTotal*: total number of retweets that a retweet receives. It was calculated using the count function in Excel to count the duplicates of the retweets in all rows and display the results in a separate column.

3.2.2.2 Independent Variables

- *TwtSntPos, TwtSntNe*: classifies a tweet with positive sentiment or negative one. It is further discussed in the data analysis section.

3.2.2.3 Control Variables

Tweet-related factors

- *TwtLength*: length of a tweet message in terms of characters. It was calculated using the LEN function in Excel to calculate how many characters are in in a cell.

- *TwtURL*: measures the number of the tweet's Universal Resource Locator (URLs). It was calculated using the COUNTIF and VLOOKUP functions in Excel to determine whether a tweet has URL(s) and counting how many, or not (0) and display the results in a separate column.

- *TwtMentd*: measures the number of mentioned usernames in a tweet, e.g. @username. It was calculated using the COUNTIF and VLOOKUP functions in Excel to determine whether a tweet has mentioned usernames and counting how many, or not (0) and display the results in a separate column.

- *Twt#*: counts the number of hashtags (#) a tweet contains. It was calculated using the COUNTIF and VLOOKUP functions in Excel to determine whether a tweet has hashtags (#) and counting how many, or not (0) and display the results in a separate column.

- *TwtMSA*: shows whether the tweet is written in Modern Standard Arabic (1), or not (0). It was classified manually by the author whose mother tongue is Arabic only for the top 100 retweets.

User-related factors

- *UsrFol*: shows a user's number of followers at the time the tweet was generated. It was retrieved automatically by Python.

Table 2 shows a summary of variables' descriptive statistics.

TABLE 2 Descriptive statistics of the dataset

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
RtwtTotal	1.008	9.862	0	1,098
TwtSntPos	0.143	0.540	0	12
TwtSntNeg	-1.469	1.534	-13	0
Twt#	0.281	1.038	0	22
TwtURL	0.524	0.516	0	3
TwtLength	98.25	42.85	1	216
TwtMentd	0.749	1.147	0	14
UsrFol	7,425	112,267	0	1.364e+07

4 HYPOTHESES TESTING

4.1 Data Analysis

4.1.1 First Dataset Analysis

4.1.1.1 Network Analysis

After collecting data from Twitter Streaming Importer plugin in Gephi, we examined the network graph diameter, betweenness centrality, closeness centrality, eccentricity, and module class. We used the Yifan-Hu layout to visualize the users' network. In order to provide clear network visualization, the Giant Component filter under Topology option was added; to clear out the nodes that are not connected to the main cluster since they tend not to contribute to the network analysis. Furthermore, a sub-filter called Degree Range with the value of 2 was added to the Giant Component filter; to filter out the nodes that have less than two connections; this made the network more manageable. The findings are discussed in the results section.

4.1.2 Second Dataset Analysis

4.1.2.1 Sentiment Analysis

Sentiment analysis in the Arabic language is limited due to the lack resources about the topic, however, the research is progressing within this area and some researchers of the field have provided a public collection of datasets that contain lexicons with already polarity-labeled (positive or negative) words, phrases, emojis, etc. in Modern Standard Arabic (MSA) and other Arabic dialects.

ElSahar and El-Beltagy (2015) built a large Arabic multi-domain lexicon for sentiment analysis; data was gathered from various website reviews with a total of 33,000 reviews on hotels, books, movies, products, and restaurants. A similar research was carried out by Aly and Atiya (2013) who classified Arabic sentiment polarity of over 63,000 book reviews and published their lexicon after. In addition, Abdulla et al. (2013) built a lexicon and made it available online; it consists of 2000 manually-labeled tweets (1000 positive tweets and 1000 negative ones) from users' opinions on different topics such as politics and arts. They also made their lexicon available online.

All lexicons were combined together in one Excel sheet which was imported after into Python. Using experiments code that has been made publicly available for scientific purposes from ElSahar and El-Beltagy (2015), the collected tweets' content from the second dataset was run through the combined lexicon to determine the sentimental polarity (polarity = positive + negative). This was done to each tweet that was collected during the data gathering period, and then the results were exported to an Excel sheet to perform further measurements for sentiment analysis (i.e. regression analysis).

4.1.2.2 Regression Analysis

To test hypotheses H2, H3, and H4, which suggest that a positive relationship exists between sentiment, tweet and user features, and retweeting quantity, the following variables were used as in previous studies; Suh et al. (2010), Stieglitz and Dang-Xuan (2013), and Zhang and Zhang (2016) in which it was demonstrated that these variables affect the retweet quantity:

- *RtwtTotal*: total number of retweets that a retweet receives.
- *TwtSntPos*: the positive sentiment of a tweet.
- *TwtSntNeg*: the negative sentiment of a tweet.
- *Twt#*: the number of hashtags in a tweet.
- *TwtURL*: the number of URLs in a tweet.
- *TwtLength*: length of a tweet message in terms of characters.
- *TwtMentd*: the number of mentioned usernames in a tweet.
- *UsrFol*: user's number of followers.

The dependent variable *RtwtTotal* represents count data which is nonnegative and integer based, with standard deviation and variance values that are larger than its mean in the second dataset, thus, the negative binomial regression analysis was applied in order to handle the over dispersion of the variables *RtwtTotal, TwtLength,* and *UsrFol*. To deal with the over dispersion issue; Stieglitz and Dang-Xuan (2013) and Zhang and Zhang (2016) log-transformed the variables before initiating the OLS regression. Similarly, a regression model was designed to interpret the effect of the variables on the dependent variable *RtwtTotal* as the following:

$$Log(RtwtTotal) = \beta 0 + \beta 1 * TwtSntPos + \beta 2 * TwtSntNeg + \beta 3 * Twt\# + \beta 4 * TwtURL + \beta 5 * Log(TwtLength) + \beta 6 * TwtMentd + \beta 7 \, Log(UsrFol) + \varepsilon.$$

Clustered robust standard errors were used in the regression analysis, because of its suitability for the data. As Zhang and Zhang (2016) stated, the clustered errors take into account that observations within groups are correlated. They added that using a large sample may lead to a significant reduction of the p value and will strengthen the significance level of the results (as in Guo et al. 2014).

4.2 Results

4.2.1 First Dataset Results

4.2.1.1 Users Network Results

Measuring the network diameter enabled us to determine the network's betweenness centrality, closeness centrality, and eccentricity values. Module class was measured after and resulted in three communities with percentages of 55.99%, 37.26%, and 6.75%. The first community, colored in red in the generated graph below (Figure 10), had 11 labeled extreme users and 9 moderate ones, which indicates that this community contains a mixture of users with different ideologies interacting with each other. The second community, with green color, had 8 moderate users and 0 extreme ones, marking it as a moderate community. The third community, with blue color, had 5 extreme users, and 0 moderate ones, hence, it is considered an extreme community. Table 3 shows the average centrality measures for each community.

TABLE 3 Network centrality measures of the sample's three communities

Centrality Measure	**Mixed**	**Moderate**	**Extreme**
Betweenness centrality	0.852	210.06	237.56
Closeness centrality	0.039	0.282	0.349
Eccentricity	0.09	2.42	3.14

From the network centrality measures, we found that users from the extreme and moderate communities have higher centrality measures than the mixed community. The extreme community has slightly higher values than the moderate community, which means that users from the extreme community have individuals with more significant influence on the network and more connectivity between users leading to more effective information dissemination. Since the *User Network* logic represents the interaction between users including retweets and mentions.#We conclude that members from the extreme and moderate communities have higher interaction rates of retweeting and mentioning other users from within their community than with the ideologically-opposed one. This confirms H1 meaning that users are more likely to respond and communicate with people who follow a similar ideology.

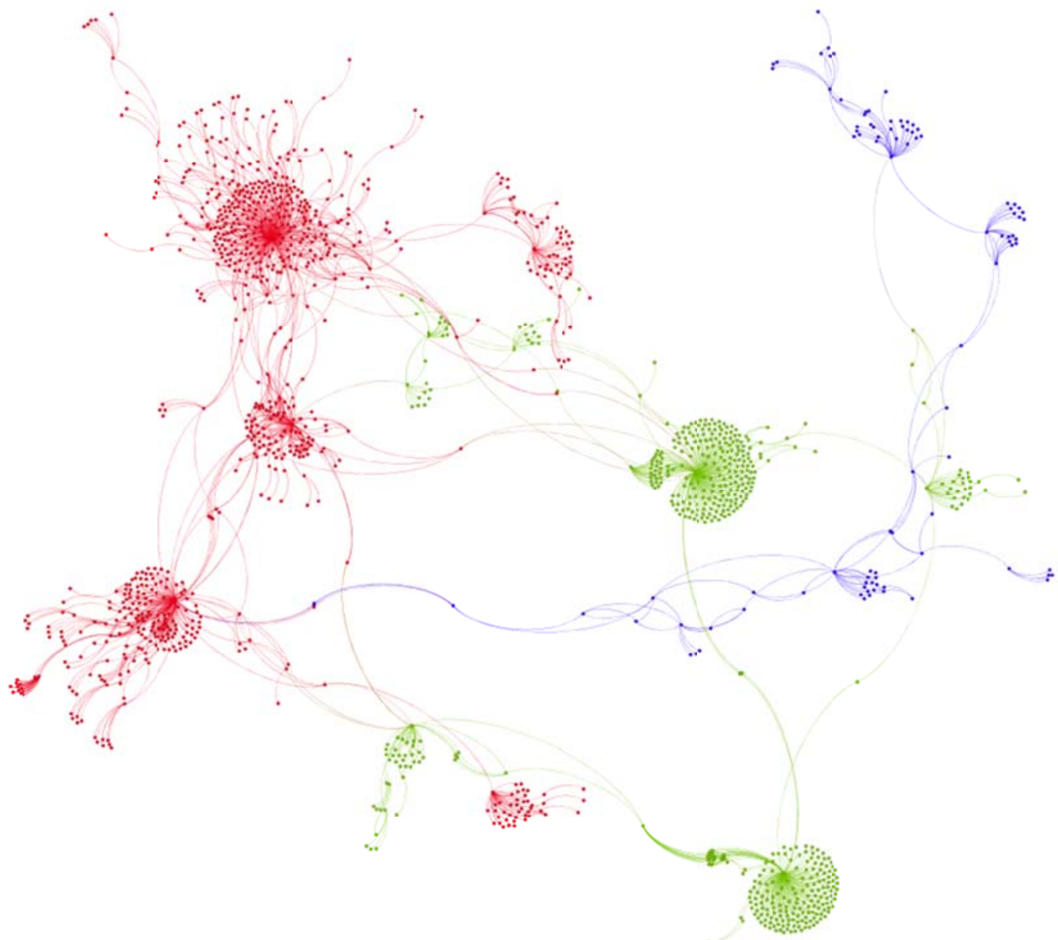

FIGURE 10 Users network graph. Nodes colored by module community (red= mixed community, green= moderate community, blue= extreme community).

For the mixed community, we suggest that the heterogeneity of the users in this community comes from similar grounds to what was described by Conover et al. (2010) as "politically motivated individuals provoke interaction by injecting partisan content into information streams whose primary audience consists of ideologically-opposed users". We believe that religiously motivated extremist and moderates act in a comparable way.

The nodes in graph 10 represent network users, and the edges represent interactions between these users (retweets and mentions). After a 4-day collection period, 1,575 nodes and 2,713 edges were created. This number was reduced to 1,436 nodes and 2,569 edges after applying the Giant Component filter to provide clear network visualization. Then, the Degree Range sub-filter was added with the value of 2 to filter out the nodes that have less than two connections which made the network more manageable ending up with 753 nodes and 1,886 edges for the final visualization of the network.

To dig deeper in the network, we used the edge network measures to examine the interaction between the extreme users and the moderate ones. The edges are directed from a node to another by retweets and mentions created by users. As more interactions are performed on the edge, it gains higher weight. Burns (2010) stated that the number of connections between two nodes determines the numerical weight of the network edge connecting them. He added that the greater weight the edge has the stronger and deeper is the relationship between users, and that the absence of highly weighted edges might indicate that the discussion is relatively free ranging and without solid and lengthy conversations between specific groups of users.

In order to investigate and compare the strength of the relationships within the extreme and the moderate groups, the network edges' weights were analyzed for the previously-labeled extreme and moderate users. Table 4 below shows the edge-weight statistics of the labeled users in the network (16 extreme users and 17 moderate ones).

TABLE 4 Edge network, weights statistics

Users	**Count**	**Ret.**	**Ment.**	**W. Mean**	**W. Std. Dev.**	**Min W.**	**Max W.**
Extreme	401	256	145	2.28	4.16	1	57
Moderate	1096	640	457	1.87	2.38	1	34

Although the moderate users had more transactions including retweets and mentions (1,096) compared to the extreme ones (only 401), the extreme users had higher edge weights on average with the value of 2.28 compared to the value of the average weights of the moderate users with 1.87. This confirms that the extreme users have stronger relationships and connectivity within their group than the moderate users.

The direction of the edges was determined in Gephi through allocating sources and targets of the edge. Meaning that, for example, if user A mentions user B, an edge from A to B is formed with user A being the source and user B being the target. Tables 5 and 6 below display the statistics of the sources and the targets transactions respectively. While the moderate users had more sources and targets transactions, still, their edges' weights were lower than the weights of the extreme edges, indicating weaker relationships among the moderate users compared to those which are formed between the extreme users.

TABLE 5 Edge network, sources statistics

Sources	**Count**	**Ret.**	**Ment.**	**Ret. W.**	**Ment. W.**	**Avg. W.**
Extreme	130	66	64	1.29	4.05	2.65
Moderate	266	124	142	1.29	3.65	2.54

TABLE 6 Edge network, targets statistics

Targets	**Count**	**Ret.**	**Ment.**	**Ret. W.**	**Ment. W.**	**Avg. W.**
Extreme	271	191	80	1.4	3.75	2.1
Moderate	830	515	315	1.33	2.17	1.65

Additionally, figures 11 and 12 present comparisons between the extreme and the moderate users by showing the interactions quantity identified by retweets and mentions, and the weights of these transactions including sources and targets own average weights.

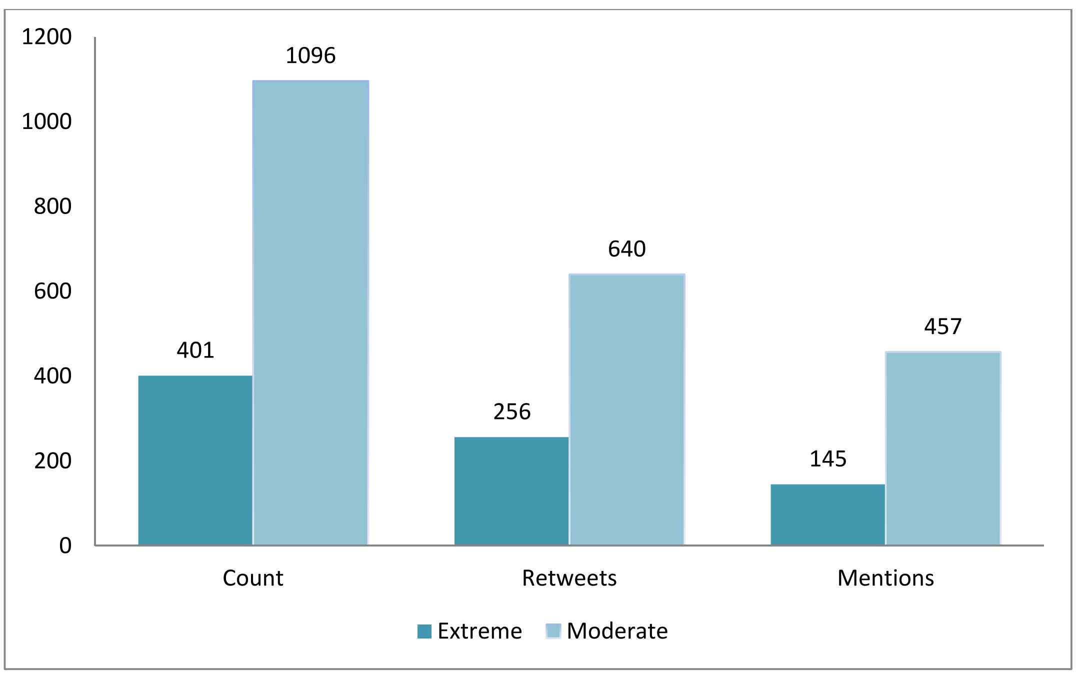

FIGURE 11 Edge network; transactions quantity of the extreme users and the moderate ones, including retweets and mentions.

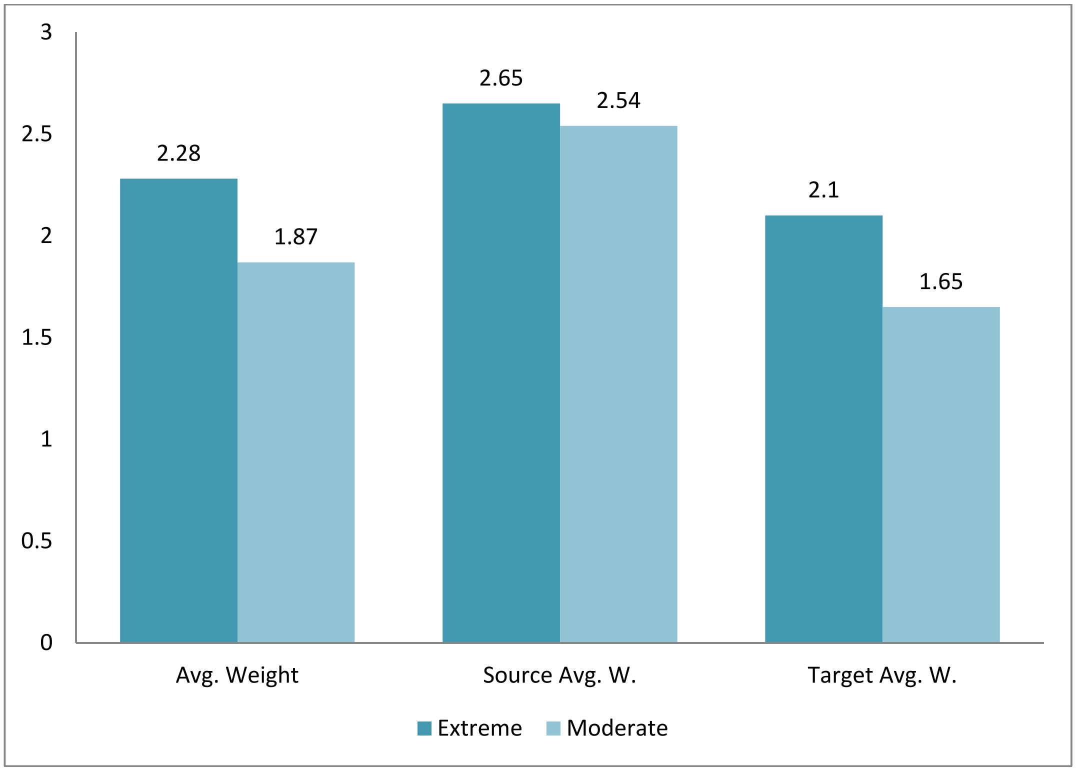

FIGURE 12 Edge-network average weights, including sources and targets of the extreme and moderate users.

4.2.2 Second Dataset Results

4.2.2.1 Regression Results

To ensure better fitting of the regression model, clustered robust standard errors were used; this made the P value more meaningful, leading the model to be statistically significant. Below in Table 7, are the overall data analysis results of the second dataset based on the regression model.

TABLE 7 Regression results

Variable	**Log(*RtwtTotal)***
TwtSntPos	-0.00694 (0.0245)
TwtSntNeg	0.0799*** (0.00863)
Twt#	0.194*** (0.0194)

TwtURL	0.144*** (0.0339)
Log(TwtLength)	0.0215*** (0.000420)
TwtMentd	-0.551*** (0.0134)
Log(UsrFol)	0.00000299*** (0.000000451)
Constant	-2.369*** (0.0367)
Lnalpha Constant	2.016*** (0.0132)
Observations	53271

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The variable *TwtSntPos* has a coefficient of -0.00694 which is not statistically significant. This means that for each unit increase on *TwtSntPos*, the expected Log(*RtwtTotal)* decreases by 0.00694, resulting in not supporting H2a . On the other hand, supporting H2b; negative sentiment in tweets has a significant positive impact on retweeting. We believe, as was shown in the findings of Baumeister et al. (2001), Rozin and Royzman (2001), and Stieglitz and Dang-Xuan (2012 & 2013) that it is the power of negativity bias on users that leads them to diffuse more negative content than neutral one.

Content features of hashtags and URLs had the most positive effect on retweeting with coefficients of 0.194 and 0.144 respectively, supporting H3a and H3b. They were followed by tweet length with coefficient of 0.0215 supporting H3c and marginally positive effect by the number of followers supporting H4. The number of mentioned names had a significant negative effect on retweeting opposing to H3d; Suh et al. (2010) explained such findings as the following "URLs and hashtags correlate positively…, whereas mentions have a negative correlation. This makes sense given that tweets are limited in the amount of content they can communicate, so having one kind of content (e.g., URLs) will tend to be exclusive of another (e.g., mentions)".

Table 8 below shows the hypotheses testing results.

TABLE 8 Hypotheses testing results

Hypothesis	**Result**
H1: *Users' interaction with ideologically-similar users has a higher frequency of retweets and mentions than their interaction with the ideologically-opposed users.*	Supported
H2a: *Tweets with positive sentiments have a higher likelihood to be retweeted comparing with tweets which are emotionally neutral.*	Not supported
H2b: *Tweets with negative sentiments have a higher likelihood to be retweeted comparing with tweets which are emotionally neutral.*	Supported
H3a: *Hashtags, have a positive effect on retweeting.*	Supported
H3b: *URLs have a positive effect on retweeting.*	Supported
H3c: *Tweet length has a positive effect on retweeting.*	Supported
H3d: *Number of mentioned names has a positive effect on retweeting.*	Opposite
H4*: Number of followers has a positive effect on retweeting*.	Supported

5 DISCUSSION

In this paper, the type of interaction between extreme Muslims and Moderate ones was examined, and their strategies to diffuse their information were investigated. Also, the effect of emotions and other tweet features on retweeting behavior was tested. The study focused on Muslims from the Arab-speaking countries because of the lack of such studies on that significant and fast-growing population with more than 370 million who have similar beliefs and speak the same language. The study contributes to research by extending literature about the Arab-Muslims' usage of Twitter and goes deeper by categorizing two groups among them; one that supports extreme approaches to diffuse and share information, and one that believes in more moderate measures to disseminate their ideology. The topic of the study was controversial because of the related sensitivity to describing a group as extreme Muslims; however, we believe that the study approach was performed ethically and followed the scientific methodology by relying on previous research and evidence to form the hypotheses and demonstrate the findings.

To test the hypotheses, two sets of data were collected; the first dataset was retrieved to examine the first hypothesis H1 and test the users' network on Twitter, it was collected from certain accounts that were classified as leaning to extremism or moderation based on their generated content. The classification of the accounts was performed by 22 Arab Muslims who have voted for the generated tweet by the account holder as extreme, moderate, or neutral. It should be mentioned that 1 generated tweet might not be enough to determine the account holder ideological leaning, but the classification process was simplified for practical reasons such as that most people would not answer the questionnaire if it was very long (the questionnaire consisted of 46 relatively-long questions), and also for the reason that it is most likely that users will generate similar content that fits their ideology; as it was described in the study for finding extremists in online social networks by Klausen et al. (2018), in which they stated that "users that engage in some form of online extremism or harassment will have very similar behavioral characteristics in social networks". We used Gephi, version 0.9.2 to analyze and visualize the first dataset, identify extreme and moderate communities, and apply the graph measures to ultimately find that users - whether they are extreme or moderate - have better connectivity and able to diffuse their information in a higher frequency of retweets and mentions within their communities rather than with the ideologically opposed users. The extreme community however, had stronger connectivity among its users, more than that within the moderate community; this was determined by the network centrality measures and supported by the network-edge weight analysis which confirmed that the extreme users have more solid relationships within their group than the moderate users. We also believe that religiously-motivated Arab-speaking Muslims act similarly to politically motivated individuals (from the study by Conover et al., 2010) in their way of diffusing their partisan content to users with opposed ideologies.

To test the hypotheses H2, H3, and H4, a second dataset was retrieved using Python, version 2.7.0, based on specific keywords that are related to Islam. 100,000 Arabic tweets were collected, which then was reduced to 53,271 tweets, each of which has generated at least one retweet. We used a Python code to detect sentiment of the retrieved tweets, the code contained a large lexicon of words, phrases, and emojis with already defined polarity (positive +1, negative -1) and worked by granting the tweet a total polarity result based on the used words in it. One of the limitations of the study is that the lexicons that were used were limited in size and domain specific. Additionally, Arabic dialects vary widely and have new terms added on a regular basis without following any rules or grammar, which makes it harder to build a highly accurate lexicon that contains all Arabic dialects. A point that was noticed during examining the tweet-related features is that the top 100 retweeted tweets of the second dataset showed that 86 of them were written in Modern Standard Arabic (MSA), which predicts the positive effect of using such language on retweeting. This might be because it is a closer type of language to Classical Arabic which is the language in the old Islamic texts and it is the literary standard language across the Arab world. Future research can focus on developing a tool that can detect MSA to measure its effect on the retweeting behavior and information dissemination.

The study findings indicate that tweets that contain an overall negative sentiment have a significant positive impact on retweeting quantity, while interestingly; positive sentiment was not statically significant to affect retweeting. Additionally, Hashtags, URLs, and tweet length have strong relationships with retweetability, and user's number of followers seems to have an effect on retweetability as well. Moreover, the number of mentioned names had a significant negative effect on retweeting. The results regarding negative sentiments are similar to findings from previous research in different fields (i.e. psychology and organizational studies) which explored negative emotions and their effect on people. Stieglitz and Dang-Xuan (2013) relied on the studies of Baumeister et al. (2001) and Rozin and Royzman (2001) to explain the Negativity Bias phenomenon; how people tend to give more weight to negative entities. We suggest that negativity bias could be one factor that could lead to such results but more detailed research is needed to support this claim. Stieglitz and Dang-Xuan, (2013) added that within a political or ideological atmosphere, people will diffuse even negative emotions if it was originally generated by someone who has similar ideological values. We believe that we have a similar case as Muslims with different backgrounds tend to retweet content even if it was negative because the original tweeter was another follower of their ideology.

The study contributes to existing literature in several fields such as information diffusion, sentiment and emotions in tweets, communication on Twitter and users' network. Although there is existing literature about the previously mentioned fields, this research offers a view on a different region, language, and religion; there are few studies on tweets generated in Arabic, and they are mostly done for political reasons such as the study by Lotan et al. (2011) and Oh et al. (2015) about the Tunisian and the Egyptian revolutions. Additionally, this study offers a view on the usage of Twitter within a religious context, which is a major topic for Muslims to diffuse information about on such a microblogging platform. The study also compares the networking behavior of two ideologically-opposed groups; extreme and moderate Muslims. This is important to grasp because it could be used as a practical implication to determine in which way these groups are communicating and recruiting more people. It could also predict the tendency of a geographical region heading towards extremism or moderation. Additionally, religious and political parties in the Arab world can use the findings to analyze what kind of speech and sentiment the crowd is more attracted to and where the influential users are located in the network. Arabic companies could also use the sentiment analysis on their products and services reviews on social media platforms and in creating advertisements that use emotions to reach more audience.

6 CONCLUSION

Nowadays, social network platforms are frequently used for information dissemination. This has affected people notions driving societies to change, and leading to form groups of members with similar ideas debating other groups with different beliefs on social-media platforms. This research examines the Arab-Muslims usage of Twitter, how they are connected and how do they disseminate their ideologies on Twitter. Different from prior studies which focus on political relevant information dissemination in the English spoken world (e.g. Conover et al. 2011), or English tweets during the Egypt Revolution (e.g. Oh et al. 2015), this study examines two ideologically-opposed groups, extreme Muslims and moderate ones, and analyzes Arabic tweets. We hope the findings of this research will provide insights regarding information dissemination in the Muslim world and help us understand polarization in the online Arab-Muslim social network communities.

REFERENCES

Abdulla, N., Mahyoub, N., Shehab, M., & Al-Ayyoub, M. (2013). Arabic sentiment analysis: Corpus-based and lexicon-based. In *Proceedings of The IEEE conference on Applied Electrical Engineering and Computing Technologies (AEECT)*.

Aly, M., & Atiya, A. (2013). Labr: A large scale arabic book reviews dataset. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (Vol. 2, pp. 494-498).

Baumeister, R.F.; Bratslavsky, E.; Finkenauer, C.; and Vohs, K.D. Bad is stronger than good. *Review of General Psychology, 5,* 4 (2001), 323–370.

Bruns, Axel. (2010). Visualising Twitter dynamics in Gephi, part 1. *Mapping Online Publics.* Retrieved from http://mappingonlinepublics.net/2010/12/30/visualising-twitter-dynamics-in-gephi-part-1/

Boyd, D., Golder, S., & Lotan, G. (2010, January). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (pp. 1-10). IEEE.

Conover, M., Ratkiewicz, J., Francisco, M. R., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. *ICWSM, 133*, 89-96

ElSahar, H., & El-Beltagy, S. R. (2015, April). Building large arabic multi-domain resources for sentiment analysis. In *International Conference on Intelligent Text Processing and Computational Linguistics* (pp. 23-34). Springer, Cham.

Forbes contributor. (2017). Twitter's surprising user growth bodes well for 2017. *#MarketMoves*. Retrieved from https://www.forbes.com/sites/greatspeculations/2017/04/27/twitters-surprising-user-growth-bodes-well-for-2017/2/#7a24c60746f8

Gephi.org. (2017). Gephi, makes graphs handy. Retrieved from https://gephi.org/features/

Global Terrorism Index. (2016). Measuring and understanding the impact of terrorism. *Institute for Economics & Peace*. Retrieved from http://economicsandpeace.org/wp-content/uploads/2016/11/Global-Terrorism-Index-2016.2.pdf

Guo, W., Straub, D., and Zhang, P. 2014. "A Sea Changein Statistics: A Reconsideration of What is Important in the Age of Big Data," *Journal of Management Analytics* (1:4), pp. 241–248.

Hoang, T. A., & Lim, E. P. (2012, May). Virality and Susceptibility in Information Diffusions. In *ICWSM*.

Honey, C., & Herring, S. C. (2009, January). Beyond microblogging: Conversation and collaboration via Twitter. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on* (pp. 1-10). IEEE.

Huda. (2017). The Importance of the Arabic Language In Islam. Why Many Muslims Strive to Learn Arabic. *ThoughtCo.* Retrieved from https://www.thoughtco.com/arabic-language-in-islam-2004035

Huda. (2017). World's Muslim population. Statistics about the Muslim population of the world. *ThoughtCo.* Retrieved from https://www.thoughtco.com/worlds-muslim-population-2004480

Islahi, A. A. (2011). Difference Between Hadith and Sunnah. *SA Rauf, Transl.), Renaissance. Retrieved August*, *13*, 2011.

Ji, X., Machiraju, R., Ritter, A., & Yen, P. Y. (2015). Examining the Distribution, Modularity, and Community Structure in Article Networks for Systematic Reviews. In *AMIA Annual Symposium Proceedings* (Vol. 2015, p. 1927). American Medical Informatics Association.

Jones, A. (Ed.). (2011). *The Koran*. Weidenfeld & Nicolson.

Kallas, Priit. (2017). Top 15 most popular social networking sites and apps [November 2017]. *DreamGrow, Your Source of Content Marketing & Social Media Information.* Retrieved from https://www.dreamgrow.com/top-15-most-popular-social-networking-sites/

Klausen, J., Marks, C. E., & Zaman, T. (2018). Finding Extremists in Online Social Networks. *Operations Research*, *66*(4), 957-976.

Levallois, Clement and Totet, Matthieu. (2017). Twitter Streaming Importer. *GitHub.* Retrieved from https://github.com/seinecle/gephi-tutorials/blob/master/src/main/asciidoc/en/plugins/twitter-streaming-importer-en.adoc

Lotan, G., Graeff, E., Ananny, M., Gaffney, D., & Pearce, I. (2011). The Arab Spring| the revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *International journal of communication*, *5*, 31.

Morey, P., & Yaqin, A. (2011). *Framing Muslims*. Harvard University Press.

Oh, O., Agrawal, M., & Rao, H. R. (2013). Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. *MIS Quarterly* (37:2), pp.407-426.

Oh, O., Eom, C., & Rao, H. R. (2015). Research Note—Role of Social Media in Social Change: An Analysis of Collective Sense Making During the 2011 Egypt Revolution. *Information Systems Research, 26*(1), 210-223.

Rozin, P., and Royzman, E.B. Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5,* 4 (2001), 296–320.

Safi, O. (2003). What is progressive Islam?. *ISIM Newsletter, 13*, 2.

Statista. (2017). Number of social media users worldwide from 2010 to 2021 (in billions). *The Statistics Portal, Statistics and Studies from more than 18,000 Sources*. Retrieved from https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems, 29*(4), 217-248.

Stieglitz, S., and Dang-Xuan, L. Impact and diffusion of sentiment in public communication on Facebook. In J. Pries-Heje and M. Chiasson (eds.), *Proceedings of 20th European Conference on Information Systems.* Atlanta: Association for Information Systems, 2012, paper 98.

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010, August). Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *Social Computing (socialcom), 2010 IEEE Second International Conference on* (pp. 177-184). IEEE.

Susarla, A., Oh, J. H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research, 23*(1), 23-41.

Toni Hirst. (2010). Getting Started With Gephi Network Visualisation App – My Facebook Network, Part III: Ego Filters and Simple Network Stats. *OUseful.Info, the blog*. Retrieved from https://blog.ouseful.info/2010/05/10/getting-started-with-gephi-network-visualisation-app-%E2%80%93-my-facebook-network-part-iii-ego-filters-and-simple-network-stats/

United Nations. (2017). World population prospects 2017. *DESA / Population Division*. Retrieved from https://esa.un.org/unpd/wpp/

Yang, J., & Counts, S. (2010). Predicting the Speed, Scale, and Range of Information Diffusion in Twitter. *ICWSM, 10*, 355-358.

Zeng, X., & Wei, L. (2013). Social ties and user content generation: Evidence from Flickr. *Information Systems Research* (24:1), pp. 71-87.

Zhang, Z., & Zhang, Y. S. (2016). How Do Explicitly Expressed Emotions Influence Interpersonal Communication and Information Dissemination? A Field Study of Emoji's Effects on Commenting and Retweeting on a Microblog Platform. In *Proceedings of the 16th Pacific Asia Conference on Information Systems*.

APPENDIX 1 QUESTIONNAIRE RESPONSES

Tweet	Extreme Votes	Moderate Votes	Neutral Votes	Decision
وجب تطبيق الحكم الشرعي بإقرار قانون لصد التبرج	20	1	1	90,9% Extreme
عروس وعروسته امام الملاء يحتضنها بحراره ويتصور معها في تحدي واضح واستفزاز للشعب المحافظ الغيور	15	4	3	68,2% Extreme
لماذا تزوجت/ي! لما لم تتزوج/ي! لماذا ترتدي هذا الزي! الخ من الأسئلة، تنتمي لفئة "التدخل فيما لايعنيك" "من حسن إسلام المرء تركه مالايعنيه"	0	17	5	77,3% Moderate
إغلاق_المحلات_وقت_الصلاة_وا جب_الأذان	16	2	4	72,7% Extreme
لابدمن #مراجعة من يوهم #الشباب انه مثقل بالذنوب حتي يعيش باقي حياته في #غم_وهم ويفقد #أغلي #نعم_الله	8	13	1	59,1% Moderate
لا تتهاونوا في سماعِ الموسيقى والغناء في مجالسكمِ وفي أفراحكم.. فإن عواقبها وخيمة وآثارها جسيمة !!	19	1	2	86,4% Extreme
يرد على مدير جامعة في دعوته للاختلاط في التعليم.. فما أحوجنا اليوم إلى هذه الردود العلمية الدامغة على دعاة الاختلاط المحرم.	16	4	2	72,7% Extreme
تهمة #إزدراء_الأديان .. هي حكم جديد بالإعدام على مشروع #تجديد_الخطاب_الديني	11	8	3	Divided
أكبر إساءة تأتي للإسلام هي من تقديس اجتهادات بعض الأشخاص وكتبهم والتي تم اعتبارها كأصلٍ في الدين	7	13	2	59,1% Moderate
سيزيد الفساد ولكن ان لقّب الجرم بالتحريم وحوسب باسم الشريعة فلن يستمر أبدا	8	6	8	Divided
كل من يغضبه إغلاق المحلات وقت الصلاة لا شك في قلبه مرض سبحان الله وهل يغضب من يرى شعائر الله تقام!	16	3	3	72,7% Extreme
قد لا توجد خطب جمعة .. وقد تكون مجرد أمور قد طرحت معبرة عن احداث خاصة في زمنهم .. واحمدوا الله على أنه لا توجد خطب والا أثقلوا على الناس وجعلوا ما جاء فيها دينًا مُلزمًا..	5	9	8	Divided
تقليد الغرب في بعض الامور ليست ثقافة إللي حصل منظر شاذ لا يناسب عاداتنا وتقليدنا ومبادئنا ك مسلمين! #عريس_يحتفل_بكورنيش_جده	8	7	7	Divided
ليس من حق المرأة ان ترتدي ماشاءت وتعبث بتعاليم رب العالمين فهنا شرع ومن لم تستحي من ربها نطلب ان تعلم الحياء بقوة السلطان	15	6	1	68,2% Extreme
النواصب شكلان: نواصب عليهم اللعنة، ونواصب لا جزاهم الله خير جزاء المحسنين	18	1	3	81,8% Extreme

سعوديات_نطلب_اسقاط_الولايه # هي اسقاط انظمه وضعيه في سائر الدوائر الحكوميه ، وليست اسقاط آباء او أهل او !..شرف او لباس	1	19	2	86,4% Moderate
هدف #الشيعة من احياء مقتل الحسين وعمل طقوس التطبير لكي يبقوا على الاحقاد والكراهية في نفوس اتباعهم ضد اهل السنة #الخوارج #الرافضة	14	5	3	63,6% Extreme
طبعاً اليوم يرقصون ... إذا كانت الدولة جادة في معالجة هذا.. #اعيدوا_هيبة_الامر_بالمعروف	18	2	2	81,8% Extreme
وعلينا مجابهة كل أصناف التطرف الفكري، على مستوى الدين والعقيدة، والتمييز العنصري، والاختلافات الطبقية، والعدل والمساواة	2	17	3	77,3% Moderate
الخرافات الفقهية والشعبية حول (السحر، الجن) مدخل لارتكاب العديد من الجرائم كالاحتيال والجرائم الجنسية	3	17	2	77,3% Moderate
هذه نتائج التحريض على التطرف من مشايخ الفتنة_ #معا_ضد_الارهاب_والفكر_الضال	1	16	5	72,7% Moderate
من أقل واجباتنا لما نشوف حساب ينشر التطرف والألفاظ البذيئة والتشكيك بأعراض الناس أن نحاربه بالبلوك	1	17	4	77,3% Moderate
هل عرفتي لماذا نهى الاسلام عن خروج المرأة متبرجة ومتزينة حتى لا تفتن الرجال ..!!	15	4	3	68,2% Extreme
ضربة جديدة لمفسرين وعلماء أحاديث التكفير والتطرف. * عن فلان عن فلان أصبحت أهم من الدين بل التفاسير عالأهواء أصبحت الدين نفسه	4	11	7	Divided
بغض النظر عن جنسية الشخص .. التطرّف ليس جنسية سعودية .. التطرّف يكمل ب عقلية الأشخاص اللي امتلئت قلوبهم حقد و كراهيه للآخرين	1	15	6	68,2% Moderate
من_علامات_المتطرف شخص # مهووس بالرّفض , يكاد أن يرفُض حتى نفسه	2	13	7	59,1% Moderate
الله ماعزنا إلا بتطبيق شرع الله وسنة رسولة والأمر بالمعروف والنهي عن المنكر وليس بنشر الرقص والغناء والاختلاط والانحلال وافساد المرأة ولذلك مهما طبل المنافقين والليبراليين لهيئة الترفيه والمنكرات	15	4	3	68,2% Extreme
لا حظوا أن المتشددين يرون أن الحرية في مجملها خاطئة، في حين أنّ المتحررين يرون أن الخطأ حرية مهما كان نوعه!! أمّا الإنسان المتزن فيرى أن الحرية حق لا يقود إلى ارتكاب الأخطاء.	2	16	4	72,7% Moderate
اذا رحتوا كوفي ومطعم وكانوا مشغلين موسيقى علموهم انه ممنوع رسمياً من الدولة واذا ما وقفوها اتصلوا على الرقم 1909 بلغوهم وراح يتجاوبون معاكم. لا تكون سبب في استمراء ونشر الحرام والمنكرات	20	1	1	90,9% Extreme
نشأت ظاهرة #شيوخ_الموضة الذين يتغيرون مثل موديلات #السيارات وعروض #الأزياء كل فترة وكل موسم تغيير جديد. هذه الظاهرة تحتاج عقابا رادعا من الدولة وكذلك دراسة نفسية لهؤلاء الشيوخ	5	11	6	Divided

عزيزي الليبرالي : هناك بلدان تسمح بالزندقة و الفجور والسكر و العهر و الشذوذ و المثلية وتستطيع أن تعيش فيها حياة بهيمية راقية	16	4	2	72,7% Extreme
اي شخص يعارض حق الغير بحرياتهم الشخصيه هو متطرف . ابى من ابى وشاء من شاء . كل من يحرض على مخالفيه هو متطرف ولو سما نفسه ليبرالي .	1	17	4	77,3% Moderate
سعوديات_نرفض_صك_الولايه # نعم نرفضها لأنها لم تكن حمايه ابدا ولم تكن تكريم كثر ماكانت استعباد وذل وقهر واهانه الكثير يعانوا من هذا الصك ومازلنا نعاني ونتمنى اسقاطها	2	18	2	81,8% Moderate
من بليات الزمان عندما يأتي(شيخ علم) ويؤصل الرذائل سباب،شتم،كذب،حقد... #تجديد_الخطاب_الديني #التسامح #التعايش	4	14	4	63,6% Moderate
مشاهدبعض النساءعلى بعض الشواطئ وفي الحدائق تقطع نياط القلوب،ذهب الخوف عندالكثيرات من الله والحياء،عبايات ضيقةرفع الملابس عند البحرللركبة	17	4	1	77,3% Extreme
الناس تصفق له وتشجعه وتزغرط له ليش ما يرمونهم باي شيء علشان عيب في مجتمعنا ولا يتكرر مررة اخرى #عريس_يحتفل_بكورنيش_جده	16	4	2	72,7% Extreme
رسالة للبراغيث|دخلنا التاريخ بلا إله إلا الله دخلنا التاريخ بنبي الملحمةﷺ دخلنا التاريخ بأبي بكر وعمر وعثمان وعلي نحن التاريخ بصحيح البخاري ومسلم	13	3	6	59,1% Extreme
أستودعوا الله دينكم البلد متجه للهاويه.. - من شاب وفتاه يرقصون بأبها -الى وافد يرقص ويضم عروسه في جده -الى ذكر يمشي بعروسه كاشفه في جده -	17	2	3	77,3% Extreme
من حق المرأه السعوديه: تكون حرة بخياراتها تعيش حياتها بالطريقه الي هي تبيها مو على مزاج وليها تتساوى حقوقها مع حقوق الذكور مثل تساوي العقوبات والضرايب الخ"على حدٍ سواء" عدم معاملتها بعنصريه على اساس فقط الجهاز التناسلي!	5	15	2	68,2% Moderate
الليبراليون، الكفار - وأذنابهم من المنافقين - يتبجحون - دائمًا - المسلمين باستخدام (العقل) ، والوصية بذلك ، ونصح غيرهم بتقديمه على (النقل) ؛ وعقولهم بين (أفخاذهم) ، وأشرفهم نزع عقاله ؛ فأبدل عقله بـ (نعاله) .	19	2	1	86,4% Extreme
تجديد_الخطاب_الديني لازم يبدأ # "و جادلهم بالتي هي أحسن"أولا ب	1	20	1	90,9% Moderate
اليتيمات السعوديات لا يخرجن نهائيا من دار الايتام، في انتظار الزوج الذي لا يأتي عادة للزواج من فتاة لا عائلة لها وكثير منهن موظفات فلما لايسمح لهن بالاستقلال في السكن هنا يتم المرأة يجعلها سجينة	2	15	5	68,2% Moderate
ماذا فعل #الازهر و #شيخ_الازهر فى #تجديد_الخطاب_الدينى لمواجهه #الفكر_المتطرف و #الارهاب ؟ #تفجير_مسجد_الروضه #سيناء #تحيا_مصر #يارب	7	11	4	Divided
الإسلام وضع #القواعد_الأساسية # وترك لنا إختيار الطريقة ليتناسب مع كل	1	18	3	81,8% Moderate

الدين لاترغمون الناس #زمان ومكان ياتجار على أفكاركم أنتم وتزعمون كذباً وبهتاناً أنها أنسنة_الخطاب_الإسلامي#مشيئة الله				
إذا سقط الحجاب فمابعده أسهل ولنا في الدول التي سبقتنا مثال وشر مثال #نطالب_بمحاسبة_سياحة_جازان #سعوديات_نفخر_بولايه_اهلنا_لنا	15	4	3	68,2% Extreme
حد اللواط ان يوضع في كيس ويقذف به من جبل شاهق او اي مرتفع شاهق ان اللواط يهز عرش الرحمن فما هذا الجلد وحكمه موجود لم نغير مالمرنا الله به	19	0	3	86,4% Extreme