

Twitter Reflections on Syrian Conflict from Turkey



Selcen Ozturkcan , İnanç Arın , Nihat Kasap , and Yücel Saygın 

Abstract The Syrian conflict is a well-known regional conflict, where Turkey is among the most affected countries in political, social, and economic terms. This study explores the Turkish public reflections on the Syria conflict by analyzing 450,000 Tweets posted in the Turkish language between Feb 1, 2015, and Feb 27, 2016. This chapter contributes to the literature by providing a broader perspective with the main research questions of: (1) What are the widely discussed topics on Twitter about Syria by the Turkish users? (2) Were these topics attracting more users to Twitter or encouraging the further engagement of the already existing users? (3) How can the fading-out characteristics of the most popular topics be described? (4) Why-under which conditions and with which features-Tweets about Syria end up extensively re-Tweeted by Turkish users? The authors report a predictive model of 86.12% accuracy to classify high and low Tweets based on the number of re-Tweets received via seven features. The analysis reveals that armed fighting, religious, and political sensitivities within the Turkish public inflate the volume of posted Tweets.

Keywords Twitter usage · Turkey · Syrian conflict · Big Data · Classification · Clustering

S. Ozturkcan (✉)

Emlyon School of Business, Écully, France

e-mail: ozturkcan@em-lyon.com; selcen.ozturkcan@lnu.se; selcen@sabanciuniv.edu

S. Ozturkcan · N. Kasap

Sabancı Business School, Sabancı University, Istanbul, Turkey

e-mail: inanc@sabanciuniv.edu

N. Kasap

e-mail: nihatk@sabanciuniv.edu

S. Ozturkcan

School of Business and Economics, Linnaeus University, Växjö, Sweden

İ. Arın · Y. Saygın

Faculty of Engineering and Natural Sciences, Sabancı University, Istanbul, Turkey

e-mail: yucel.saygin@sabanciuniv.edu

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

B. Alareeni and A. Hamdan (eds.), *Financial Technology (FinTech), Entrepreneurship, and Business Development*, Lecture Notes in Networks and Systems 486,

https://doi.org/10.1007/978-3-031-08087-6_58

1 Introduction

The Syrian conflict led to a major impact on many aspects of life, from individual to public, in many parts of the globe (Issa 2016). Longer than World War II, the war in Syria has led to widespread profound global effects, with many leaving their homes to seek safety elsewhere in Syria or abroad. According to the UN refugee agency (UNHCR), over 5.5 million individuals moved abroad, in addition to the 6.5 million that have been internally displaced within Syria's borders (Aljazeera Homepage 2021). UNHCR (2019) indicates that the 1951 Geneva Convention Relating to the Status of Refugees describes a refugee as any person who, "owing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country".

Reports indicate that the economic consequences exceed \$35 billion, with more than 5 million people relocated (Connor 2018). Among Syria's immediate neighboring countries, Turkey hosted the largest population of Syrian refugees (Orhan 2015) that consisted of 3.4 million (Connor 2018) individuals. Some reports on refugee studies indicate emerging tensions among host communities, displaced Syrians, and humanitarian policymakers and practitioners (Chatty 2015; Phillips 2012), while others underline the positive impact of Syrian businesses in Turkey (Karasapan 2016). There is an obvious impact on Turkish society with a population of 80 million as 2.7 million Syrian refugees (UNHCR 2016) are now registered in Turkey. In this respect, this chapter aims to utilize Twitter analysis to explore the reflections present in the Turkish social networks about Syria and the related sub-topics. The findings are based on the analysis of 450,000 Tweets that were streamed between Feb 1, 2015, and Feb 27, 2016, with the content in the Turkish language. The main research questions focused on this chapter are as follows: (1) what are the widely discussed topics on Twitter about Syria by the Turkish users, (2) were these topics attracting more users to Twitter, or encouraging the further engagement of the already existing users, (3) how are the fading out characteristics of the most popular topics described, (4) why-under which conditions and with which features-Tweets about Syria end up extensively Re-Tweeted by the Turkish users?

The remainder of this chapter is structured in the following manner. The next sections provide an overview of existing literature as well as insight into the research context. Then, data and methodology are described, comprehensive analysis and results are provided, and significant findings are discussed. Finally, the chapter is concluded by offering perspectives relevant to further research.

2 Background

Twitter, a micro-blogging social media, which was initially adopted for networking and entertainment (Howard 2008; Rui et al. 2013) is also used in cases of social

and political movements (Della Porta 2014; García-Martín and García-Sánchez 2015; Gerbaudo 2012; Molaei 2015; Penney and Dadas 2014; Starbird and Palen 2011; Stieglitz and Dang-Xuan 2013; Theocharis et al. 2014; Tremayne 2013). As a cyber-*ekklesiá*¹ with effortless accessibility and easy information sharing, its users continuously and contagiously declare their emotions and ideas on topics of their choice (Berger and Milkman 2012; Peters et al. 2013; Shi et al. 2014; Subramani and Rajagopalan 2003). With its 320 active million users worldwide (Smith 2016), Twitter produces immense amounts of data to probe temporal behavioral patterns (Jacobs 2009; Kourouthanassis et al. 2015; Metallo and Agrifoglio 2015). Twitter data analysis uncovers meaningful findings on individual and group Tweeting characteristics and often reveals situational phenomena while predicting future events (Jacobs 2009; Savage 2011). Among Twitter's top markets of active users, Turkey ranked 8th with a share of 3.0% of global users (Richter 2013). Twitter was the 7th most popular website in Turkey (Alexa Homepage 2020), with more than 80% of the population in Turkey having an account (Akar and Dalgic 2018). Twitter was considered a vast data source, allowing research to capture and analyze the public's reflections in Turkey, particularly in under-covered news surrounding social phenomena. For example, previous literature reported wide use of Twitter in Turkey during the Gezi Park social protests (Ogan and Varol 2017; Ozturkcan et al. 2017).

Twitter analysis has been used on a wide array of research topics, which also includes political science. For example, the German federal election results had been predicted successfully by Twitter analysis (Tumasjan et al. 2010). The study of Tweets is also used in understanding the spread of phenomena concerning conflict resolution and emergency management. For instance, Twitter exhibits cues that can enable analyzers to detect real-time earthquake events, specifically an emergency (Sakaki et al. 2010, 2013).

A well-known regional conflict with recent worldwide effects is the Syrian conflict, which is cited as the source of "one of the largest movement of migrants and refugees from Asian, African and Middle East countries towards Europe" (Coletto et al. 2016). Turkey is the most affected country of the Syrian conflict in political, social, and economic terms and is living a dramatic demographic change as one of the main host countries of Syrian refugees. In Turkey, 80% of Syrian refugees live in urban areas, while some camps are built for their accommodation (Gabiam 2016). Aras and Mencutek (2015) showed how variations in foreign policies often immediately reflect upon the different states' responses as well as changes to a particular state's approach. Turkey first adopted an open-door policy towards Syrian refugees welcoming all; however, upon confrontation with massive flows of refugees, it had to revise the foreign policy orientations. Heisbourg (2015) claims that one of the quickest and least complicated policies for the EU to implement is financial support to the UN High Commissioner for Refugees in the region to cope with refugee flows from Syria. The research also suggests that the negotiation between the EU and

¹ A political assembly of citizens of ancient Greek states; especially: the periodic meeting of the Athenian citizens for conducting public business and for considering affairs proposed by the council (<http://www.merriam-webster.com/dictionary/ecclesia>).

Turkey for funding to support refugee relief in Turkey is vital to reducing refugee flows since Turkey has become a significant transit point for refugees and other migrants trying to reach Europe (Gabiam 2016). One of the main issues regarding the Syrian crisis is the education of refugee children who need access to primary education at all levels. Bircan and Sunata (2015) mention that collaboration among public and private partners at the local, national and international levels is crucial for the education program development mainly due to lack of financial matters. All in all, as reported in past research public's perceptions of refugees and migrants, such as concerns with security issues, beliefs that migrant workers might pose a threat to public safety, carry diseases, and compete for jobs and national resources with the natives can easily lead to misperceptions that would hinder the integration of refugees and migrants (Sunpuwan and Niyomsilpa 2012).

Dekker and Engbersen (2014) argue that social media provides new communication channels in migration networks and facilitates migration. Social media helps people to access widespread informal information and thus expands the perspectives of candidate migrants. Refugees perceive mobile phones as a vital tool to enable social and economic networks to remain strong. Wall et al. (2015) find that refugees view the phone within a broader political context. It was because of mobile phone use that the world noticed what was happening in Syria. In this regard, Tyshchuk et al. (2014) provide an essential insight into how the international community built an opinion about the Syrian gas attack via available information on Twitter while few journalists and international observers were available on the ground to offer any external validity checks. In addition, refugees are often able to supplement and reinforce professional news with their networks. Maitland and Xu (2015) show that limited capacity in the cellular infrastructure and an over-dependence on mobile data might create downfalls for refugees as they often access the Internet solely via their mobile phones.

The emerging field of research on uses of Social Media during social movements, crises, and conflicts includes several findings with regard to digital activism (Jansen 2010; Leuven et al. 2015; Sandoval-Almazan and Gil-Garcia 2014; Tufekci 2013), as well as the use of the visual propaganda in international conflicts (Seo 2014). However, most of this past research focuses on a narrow group of people since digital activists are individuals who engage online for social change. There is a lack of research in understanding the reflections of a broader population other than those calling for a social change in social media. This chapter provides a broader perspective of the host population in Turkey in researching reflections of the population on the Syrian conflict by analyzing Twitter reflections from Turkey. Following the work of Coletto et al. (2016) on addressing European level questions in terms of populations' perception of the refugee crisis, similar country-level questions are tackled: "How is the population in perceiving the refugee crisis phenomenon? What is the general opinion? How do events influence perception? What is the impact on public opinion of news related to refugees? How does perception evolve in time?" specifically for Turkey.

3 Data and Methodology

A Twitter Stream API², which allows for a small random sample of all related Tweets, is employed to collect real-time publicly available Tweet contents and features that contain the keyword of “Syria” (“Suriye” in Turkish) between Feb 1, 2015, and Feb 27, 2016, on a 24-h basis. There was no specific reason to choose these days other than ensuring a term longer than a year-long period for better reflective analysis. A total of 450 000 Tweets were collected, all in the Turkish language, to reflect upon how the host country’s members view the refugees and the refugee conflict. Unfortunately, as a minor limitation, the server employed received an attack and lost its data connection between Mar 24, 2015, and May 8, 2015. The used methodology is built on a social media analytics framework (Fan and Gordon 2014) that has incorporated other similar social phenomena revealing research (Çevik et al. 2015; Ozturkcan et al. 2019).

The distribution of daily Tweet numbers, daily user numbers, and cumulative user numbers are explored to comprehend the data characteristics concerning research questions 1 and 2. Additionally, the most re-Tweeted Tweets that provided the highest spread and their related dispersion are analyzed. The acceleration of spread that the highly re-Tweeted Tweets amass until reaching a saturation point is investigated regarding the 3rd research question. Here, the saturation point is defined as when a specific highly re-Tweeted Tweet lost its acceleration but started receiving lower numbers of re-Tweets. For this purpose, the top highest re-Tweeted 50 Tweets were extracted from the data set and analyzed their spread and dispersion. Next, a deeper analysis based on the Tweet’s content characteristics is conducted by increasing the number of Tweets included in the analysis step by step. Some content-based features are extracted to construct a learning model by using data mining techniques. Lastly, the prediction capability and accuracy of re-Tweet acceleration of our model is tested with 3 000 Tweets.

4 Analysis and Results

As mentioned earlier, the chapter focused on some main research questions, which were as follows: (1) what are the widely discussed topics on Twitter about Syria by the Turkish users, (2) were these topics attracting more users to Twitter, or encouraging the further engagement of the already existing users, (3) how are the fading out characteristics of the most popular topics described, (4) why, under which conditions and which features of Tweets about Syria end up extensively Re-Tweeted by the Turkish users?

The following sections of this chapter present the relevant results in this same respective order.

² Twitter API Rate Limits from Twitter Developers: <https://dev.twitter.com/rest/public/rate-limiting> Retrieved on May 12, 2016.

4.1 *Reflection of Real-Life Phenomenon in Tweets*

Previous research indicates that Tweet volume is indicative of social trends taking place in real life (Molaei 2015; Penney and Dadas 2014; Theocharis et al. 2014; Tremayne 2013). Twitter, in particular, is known as a powerful computer-based interactions medium that minimizes the gap between the virtual and the real world (Capece and Costa 2013). Therefore, the chapter analyzed the collected data to understand the distribution of the total Tweets per day. Ten days had more than 5 000 Tweets posted (Table 1). The events were mapped with relevant news concerning the top 10 highest numbers of Tweets posted daily to the Twitter volume time-series (see Fig. 1). The peak day in the data set was on Jul 20, 2015, with 11,722 Tweets posted. Common themes that lead to an increase in the number of Tweets posted are identified by chronologically associating real-life events and news. These themes were revealed by the consensus of researchers involved in the study by observing an inter-rater reliability score of 0.7. When the events were categorized with the relevant news that occurred on specific peak dates (Table 1), a common theme emerged around the armed conflict concept. Religious and political sensitivities and in-group collectivistic cultural traits among the Turkish public were among the other concepts that were identified. Turkish culture is often described as high on power distance and in-group collectivism (GLOBE 2016; Hofstede 1980), while it is deemed to carry paternalistic values (Arcan 2001; Kabasakal and Bodur 1998; Wyer et al. 2009). House et al. (2004, p. 30) defined in-group collectivism as “the degree to which individuals express pride, loyalty, and cohesiveness in their organizations or families”. The high volume of Tweets posted upon news about border issues (249 people were arrested at the border; ‘wall’ to protect the Syrian border; Cilvegozu border gate was closed), and beyond border action (Humanitarian aid was delivered to the enclaves; land operations to Syria is necessary) could be related with such in-group collectivism and paternalism aspects of the Turkish culture. On another note, it is worth considering that legal action towards Tweets due to various reasons in the recent past (Watters and Ziegler 2016) could have led to some level of self-censoring on posting Tweets, too.

As detailed in Table 1, the top 10 highest of Tweets were recorded in rows 1 to 10, with armed conflict observable in the daily events corresponding to each of these dates. News on Assad’s attack on the Turkmen village and Suruc attack, together with ISIS’s use of chemical weapons in Syria and Iraq on Jul 19, 2015, was reflected with the highest number of Tweets ($N = 11\,722$) in the dataset on Jul 20, 2015, by tapping into all three categories of armed conflict, religious and political sensitivities, and in-group collectivistic cultural traits. A similar pattern of capturing all three categories was observed on Jun 27, 2015, Sep 10, 2015, Feb 15, 2016, and Feb 14, 2016, in the descending order of captured number of Tweets. Also, Jan 12, 2016, and Jun 28, 2015, were the two dates where only the armed conflict was observed in the Tweets. This prevalence of the armed conflict category may be due to significant major events that took place on these dates. On Jan 12, 2016, a massive explosion took place at

Table 1 Top 10 highest numbers of Tweet posting days and related real-life events

	Date	Number of tweets	Real-Life occasion	Category of the event (1 = armed conflict; 2 = religious and political sensitivities; 3 = in-group collectivistic cultural traits)
1	Jul 20, 2015	11 722	Assad attacked the Turkmen village; Suruc attack; (Jul 19, 2015: ISIS used chemical weapons in Syria and Iraq)	1, 2, 3
2	Jan 12, 2016	11 146	An explosion occurred in Sultanahmet; Operation to ISIS in 5 cities	1
3	Jun 28, 2015	10 163	Sound of cannons across the Kilis; 249 people was arrested at the border;	1
4	Jun 27, 2015	9 805	ISIS attacked Al-Hasakah; The second largest massacre in Kobani; The 'wall' protection to Syrian border; Kobani again in YPG's hands	1, 2, 3
5	Feb 22, 2015	6 533	The place of Suleyman Shah Tomb has been changed; Tomb was evacuated, conflicts were increased	1, 2
6	Jul 23, 2015	6 248	Conflict in Kilis, 1 Turkish soldier was killed; Turkish F-16 hit the ISIS targets; Early elections commentary from the leader of the main opposition party	1, 2
7	Feb 18, 2016	6 014	Humanitarian aid was delivered to the enclaves; the Prime minister unclused the perpetrators	1, 3
8	Sep 10, 2015	5 985	Cilvegozu border gate was closed; 5.5 tons of bomb in trucks loaded with onions	1, 2, 3
9	Feb 15, 2016	5 467	Opened fire from YPG region in Syria, how soldiers began to respond; 50 kg of explosives were seized in Şanlurfa; YPG took control of 70% of Tel Rifat	1, 2, 3
10	Feb 14, 2016	5 309	Minister Yilmaz said TSK is hitting YPG positions; A concrete wall is put up at the Turkey-Syria border; land operations to Syria is necessary	1, 2, 3

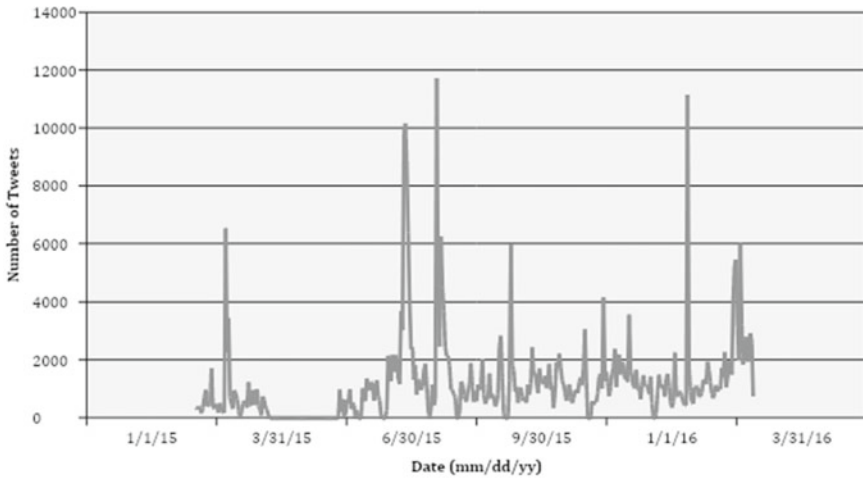


Fig. 1 Daily number of Tweets (Feb 01, 2015–Feb 27, 2016)

the touristic and historic Sultanahmet district of Istanbul in the former. On Jun 28, 2015, some 249 individuals were arrested at the Turkish border following the sound of cannons across Kilis, a Turkish town at the border.

4.2 *Attracting New Users to Post*

There is a similar pattern between the daily number of Tweets and the daily number of distinct users in the dataset (see Fig. 1 and Fig. 2), which might seem confusing and misleading that these two charts are the same. We conclude that users, who prefer to post Tweets about Syria, choose to post only 1–2 Tweets per day. The high resemblance in the patterns of distinct daily users and Tweet numbers' distributions indicates that a wide range of individuals chose to reflect on the topic of Syria. The cumulative numbers of users increased day by day, reaching more than 175,000 users towards the end of our data set. Therefore, it is concluded that Tweets containing Syria have not only been posted by a small specific group of people but indeed attracted new individuals to post (see Fig. 3).

4.3 *Spread and Fade Out Characteristics*

The highly re-Tweeted top 50 Tweets are identified (Table 2). Tweet1, which was posted by user0 on Jul 20, 2015, 12:26:43, received the highest number of re-Tweets

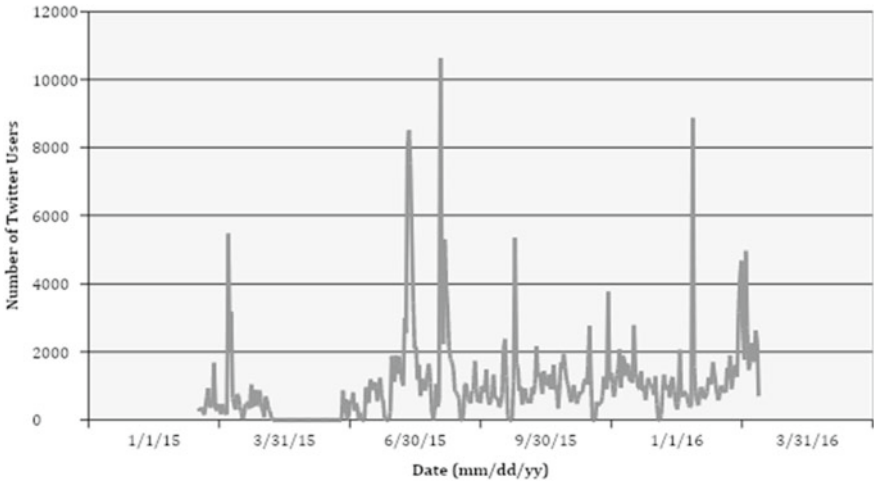


Fig. 2 Daily number of Twitter users (Feb 01, 2015–Feb 27, 2016)

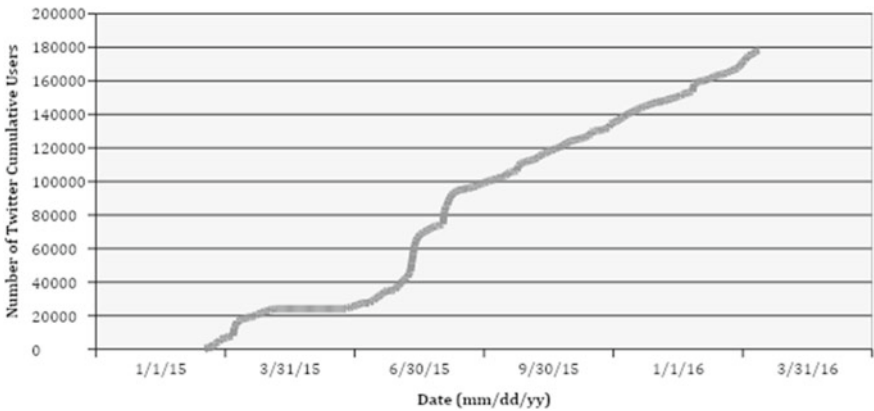


Fig. 3 Daily cumulative numbers of Twitter users (Feb 01, 2015–Feb 27, 2016)

with a total of 9,973. User0 had 1,873,272 followers when posting Tweet1. The re-Tweet line has exercised a slope of 305.64 before reaching a saturation point.

To answer the third research question (“how are the fading out characteristics of the most popular topics could be described”), the subsequent analysis focused on the characteristics of the most popular Tweets while reaching their saturation point. It is observed that some Tweets experienced a sharp decline in re-Tweets once a saturation point was reached. On the other hand, some Tweets enjoyed a relatively smooth fall with a lower diminishment of re-Tweets. The sharpness of re-Tweet decline, which is an indicator of Tweets fade-out characteristics, is examined for the top 50 re-Tweeted Tweets (Table 2). Tweet6 exercised the sharpest slope (706.50)

Table 2 Details of highly re-Tweeted top 50 Tweets and posting accounts

Tweet ID	TwitterUser ID	Number of followers	Timestamp of Tweet	Number of re-Tweets accumulated	Slope of the re-Tweet trendline
Tweet1	user0	1 873 272	Mon 20/07/15 12:26:43	9 973	305.64
Tweet2	user1	1 277 607	Fri 26/06/15 23:56:12	4 908	188.45
Tweet3	user2	6 614 251	Sat 27/06/15 21:52:22	3 851	109.41
Tweet4	user1	2 610 948	Sun 14/02/16 22:52:37	1 633	61.62
Tweet5	user3	37 275	Thu 12/02/15 20:25:58	1 530	461.50
Tweet6	user4	16 810	Sun 22/02/15 12:02:36	1 481	706.50
Tweet7	user5	73 650	Thu 25/02/16 11:51:55	1 095	373.70
Tweet8	user6	640 376	Thu 13/08/15 15:20:54	1 070	106.38
Tweet9	user7	10 995	Mon 23/02/15 21:42:16	1 053	524.00
Tweet10	user8	1 520 728	Sun 03/01/16 13:14:29	1 020	151.17
Tweet11	user8	1 521 315	Tue 12/01/16 11:39:20	988	111.61
Tweet12	user9	884 047	Mon 29/06/15 15:52:58	765	46.84
Tweet13	user10	33 886	Tue 12/01/16 14:04:41	730	110.57
Tweet14	user11	47 849	Thu 29/10/15 04:18:42	726	139.20
Tweet15	user12	381 451	Fri 25/09/15 08:25:12	724	91.57
Tweet16	user11	58 074	Sun 14/02/16 10:40:03	722	35.74
Tweet17	user13	634 235	Fri 24/07/15 07:52:51	712	133.63
Tweet18	user14	19 260	Thu 10/09/15 15:10:10	674	137.80
Tweet19	user15	676 881	Wed 22/07/15 18:14:58	670	96.04
Tweet20	user16	2 747	Sun 08/02/15 20:25:54	645	322.50
Tweet21	user17	50 234	Sun 28/06/15 12:30:48	644	201.60

(continued)

Table 2 (continued)

Tweet ID	TwitterUser ID	Number of followers	Timestamp of Tweet	Number of re-Tweets accumulated	Slope of the re-Tweet trendline
Tweet22	user18	166 810	Sun 22/11/15 09:24:10	641	98.89
Tweet23	user19	316 294	Mon 20/07/15 11:41:33	609	186.20
Tweet24	user20	190 091	Mon 20/07/15 18:39:19	600	102.86
Tweet25	user21	23 553	Thu 23/07/15 14:43:11	590	125.30
Tweet26	user22	535 194	Sun 14/06/15 21:39:45	568	28.57
Tweet27	user23	446 603	Mon 29/06/15 18:22:27	559	73.14
Tweet28	user24	166 660	Sun 28/06/15 06:38:20	530	56.06
Tweet29	user25	111 404	Thu 09/06/15 21:35:27	529	86.09
Tweet30	user26	2 345	Wed 01/07/15 17:32:13	517	63.35
Tweet31	user27	85 476	Sun 22/02/15 14:51:54	515	36.59
Tweet32	user8	1 521 326	Tue 12/01/16 14:27:35	511	74.03
Tweet33	user28	2 520 567	Sun 22/02/15 05:48:26	503	35.91
Tweet34	user29	84 693	Wed 08/07/15 09:51:00	485	81.34
Tweet35	user11	50 877	Fri 13/11/15 22:33:04	474	26.69
Tweet36	user30	300 507	Sun 28/06/15 22:23:51	466	51.50
Tweet37	user31	1 556 018	Sun 22/02/15 10:28:58	445	90.50
Tweet38	user32	44 427	Thu 02/07/15 19:49:44	442	25.83
Tweet39	user33	264 823	Wed 17/02/16 23:55:25	439	218.00
Tweet40	user34	40 102	Sat 28/11/15 00:11:09	439	26.74
Tweet41	user35	57 454	Wed 24/02/16 19:01:11	432	108.90
Tweet42	user6	694 295	Mon 18/01/16 10:15:28	424	41.88

(continued)

Table 2 (continued)

Tweet ID	TwitterUser ID	Number of followers	Timestamp of Tweet	Number of re-Tweets accumulated	Slope of the re-Tweet trendline
Tweet43	user36	62 246	Sun 22/02/15 10:26:10	418	209.00
Tweet44	user37	17 110	Thu 23/07/15 14:51:34	412	101.70
Tweet45	user38	4 220	Tue 12/01/16 11:51:05	412	125.00
Tweet46	user39	106 591	Thu 10/09/15 14:37:51	410	204.00
Tweet47	user40	302 911	Sun 15/02/15 16:14:51	407	62.94
Tweet48	user41	2 987 212	Fri 25/12/15 23:55:36	404	31.72
Tweet49	user42	96 658	Wed 22/07/15 07:32:08	402	401.00
Tweet50	user41	2 686 033	Tue 30/06/15 01:07:47	400	58.74

while amassing 1,481 re-Tweets. On the other hand, Tweet3 reached some 3,851 re-Tweets but observed a smoother trend with the most curved slope (109.41) with the highest number of followers. Tweet38 followed the smoothest slope with only 442 re-Tweets (see Fig. 4).

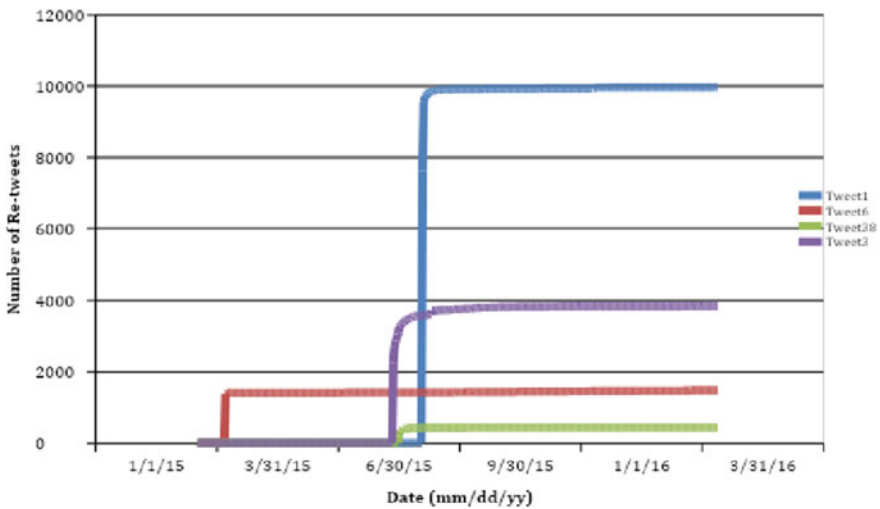


Fig. 4 A sample spread and fade-out patterns of some highly re-Tweeted Tweets

Further analysis was conducted with a broader sample containing the top 500 most re-Tweeted Tweets in the next section.

4.4 Conditions and Features Leading to High Re-tweet

It is possible to extract different characteristics from a Tweet. For instance, user-based features can be obtained, such as the user's follower number or the user's location. Time-based features can be extracted from a Tweet as well. On the other hand, analysis of the text written in the Tweet often referred to as the content-based analysis, is also possible. In this regard, some class-specific words (tokens) following the Information Retrieval Technique introduced by Rajaraman and Ullman (2011) were extracted.

The number of re-Tweets was checked initially, then the average re-Tweet level and the average re-Tweet level \pm variance were calculated. Then, re-Tweets at a level more than the "average of re-Tweets + the variance" were classified as "high"; while those less than the "average of re-Tweets-the variance" were classified as "low" class. Lastly, "high" classified re-Tweets were further examined to understand any further possibilities to identify sub-levels. These sub-levels appeared around 500 and 5,000 re-Tweets, which were incorporated in the further analysis detailed below.

Regular attributes of a learning model are identified based on the content features of the highly re-Tweeted top 500 Tweets. Propagation velocities of the Tweets are extracted via the Simple Linear Regression method to discover relations in the data model. The Tweet distribution data for each Tweet is fit into a line using Simple Linear Regression; then, their slopes are calculated and categorized as high and low according to the normal distribution ($\mu = 76.08$; $\sigma = 69.11$). Tweets with a slope higher than $\mu + \sigma$ are categorized as the high class ($n = 46$) and the remaining Tweets as the low class. For the fourth research question ("why, under which conditions and which features of Tweets about Syria end up extensively Re-Tweeted by the Turkish users"), six regular content-based and user-based features from the Tweets, namely containsHashtag, containsLink, numberOfFollowers, numberOfCapitalLetters, containsPoliticWords, and numberOfTotalRTs were extracted. Their detailed content features are also provided in Table 3.

In the rest of this section, learning models are composed using the features mentioned above and some other features, which will be mentioned in the following parts to make predictions regarding the fourth research question. RapidMiner, an open-source Data Mining tool, is used in stratified k-fold cross-validation ($k = 10$) to build the model and estimate the accuracy of the model (Table 4).

Results indicated well overall accuracy (90.60%) but low recall value for the high class. Naïve Bayes classifier was more likely to make predictions on low since most of the records ($n = 454/500$) belonged to the low class. Feature subsets are then selected via the brute force feature selection technique to identify the best feature combination for the developed classifier. Although accuracy improved (Table 4) upon

Table 3 Attributes of the learning model

Attribute	Content-feature
<i>containsHashtag</i>	It is a binomial attribute. It is <i>true</i> for the Tweet that contains a hashtag; <i>false</i> otherwise
<i>containsLink</i>	It is a binomial attribute. It is <i>true</i> for the Tweet that contains a link; <i>false</i> otherwise
<i>numberOfFollowers</i>	The follower number of the Twitter user at the time when the Tweet had been posted
<i>numberOfCapitalLetters</i>	The number of capital letters in the Tweet text
<i>containsPoliticWords</i>	It is a binomial attribute. We have generated a list of political words, consisting of political party names, politician names, some political keywords, etc. This attribute is <i>true</i> if the Tweet contains any of these political words; <i>false</i> otherwise
<i>numberOfTotalRTs</i>	It reflects how many times a particular Tweet has been re-Tweeted

Table 4 Confusion matrix on predicting low and high labels

	True high	True low	Class precision
Predicted high	5	6	45.45%
Predicted low	41	448	91.62%
Predicted total	46	454	
Class recall	10.87%	98.68%	
Accuracy: 90.60 ± 2.84% (micro: 90.60%)			

this feature selection operation, for prediction purposes, *numberOfTotalRTs* is to be regarded as a unique attribute similar to the slope.

Acceleration Analysis. The types of dispersion that Tweets amass over time vary. Typically, high influencer Tweets receive more elevated numbers of re-Tweets in an extended period. However, the acceleration of re-Tweet accumulation—whether the re-Tweets accumulated slowly or quickly, also provides information about the nature of their spread. Therefore, an acceleration analysis is conducted to understand the nature of Tweet spread better.

In addition to classifying *numberOfTotalRTs* as a special attribute, another special attribute, *numberOfDaysForSaturation*, which indicated the number of days it takes for a given Tweet to reach the ending of all its re-Tweets, is included in the analysis. Therefore, the analysis is conducted with three labels (slope, *numberOfTotalRTs*, and *numberOfDaysForSaturation*) and five regular attributes that were used as the feature(s) for learning (*containsHashtag*, *containsLink*, *numberOfFollowers*, *containsPoliticWords*, and *numberOfCapitalLetters*).

There was a major difference in data size in the low (454 Tweets) and the high (46 Tweets) classes. To remedy this difference, the training set was increased from 500 to the top 1,000 highest re-Tweeted Tweets. This ensured an improved balance between the classes. Results are provided below for slope, *numberOfTotalRTs* and *numberOfDaysForSaturation*:

Slope. Following the 68-95-99.7% Rule, 91 high and 909 low values were detected. Since unbalanced class sizes were obtained, under-sampling, which uses only a subset of the majority class in a very efficient way, was applied (Liu et al. 2008). Therefore, the analysis selected 91 low Tweets and used all 91 high Tweets to experiment on a balanced dataset. Cross-validation with two classifiers using the five regular attributes, namely containsHashtag, containsLink, numberOfFollowers, numberOfCapitalLetters, and containsPoliticWords, were used for Naïve Bayes and K-NN (k = 22) analysis (Table 5). Results indicated that K-NN performed better than Naïve Bayes, with 66.52% accuracy.

numberOfTotalRTs. The 26 high and the 26 low Tweets were used for adopting the 68-95-99.7% Rule. Again, there was an unbalanced dataset situation (26 high, 974 low Tweets); therefore, the under-sampling method was applied to balance class sizes. Results indicated that both K-NN (k = 22) and Naïve Bayes performed with equal accuracy at the 72.00% level (Table 6).

numberOfDaysForSaturation. 134 high and 134 low Tweets were used for adopting the 68-95-99.7% Rule. K-NN performed slightly better than Naïve Bayes, with an almost 4% overall accuracy difference. The k value for the K-NN classifier that produced the best accuracy results is given in Table 7.

Next, cluster analysis was employed to separate data into two different groups with the aid of a similarity measure in between them.

Cluster Analysis. Cluster analysis was used to group the specified class labels (high and low) for special attributes, as it allowed for two different clusters via calculation of

Table 5 Confusion matrix on predicting low and high labels for slope

	Naïve Bayes			K-NN (k = 22)		
	True high	True low	Class precision	True high	True low	Class precision
Predicted high	39	31	55.71%	61	31	66.30%
Predicted low	52	60	53.57%	30	60	66.67%
Predicted total	91	91		91	91	
Class recall	42.86%	65.93%		67.03%	65.93%	
Accuracy	54.36 ± 8.97% (micro: 54.40%)			66.52 ± 8.53% (micro: 66.48%)		

Table 6 Confusion matrix on predicting low and high labels for numberOfTotalRTs

	Naïve Bayes			K-NN (k = 22)		
	True high	True low	Class precision	True high	True low	Class precision
Predicted high	14	2	87.50%	12	0	100.00%
Predicted low	12	24	66.67%	14	26	65.00%
Predicted total	26	26		26	26	
Class recall	53.85%	92.31%		46.15%	100.00%	
Accuracy	72.00 ± 18.33% (micro: 73.08%)			72.00 ± 16.00% (micro: 73.08%)		

Table 7 Confusion matrix on predicting low and high labels for numberOfDayForSaturation

	Naïve Bayes			K-NN (k = 50)		
	True high	True low	Class precision	True high	True low	Class precision
Predicted high	25	11	69.44%	69	44	61.06%
Predicted low	109	123	53.02%	65	90	58.06%
Predicted total	134	134		134	134	
Class recall	18.66%	91.79%		51.49%	67.16%	
Accuracy	55.24 ± 5.07% (micro: 73.08%)			59.27 ± 5.54% (micro: 59.33%)		

a similarity measure between them. Dispersion of each Tweet is previously presented by the relevant best fitting line(s) obtained via Linear Regression, where the best fitting line is characteristically deemed representative of the Tweet that it is obtained from. Singhal (2001) stated that the angle between any two vectors could be used to measure similarity among them. Particularly, the cosine of the angle can be used for identifying the numeric similarity (1.0 = identical vectors; 0.0 = orthogonal vectors). Similarly, vector clusters were obtained by adopting the cosine similarity metric. The cluster of Tweets, represented by best fitting line vectors, has been examined by two different clustering techniques, namely the agglomerative clustering and the k-means clustering. In agglomerative clustering, a type of hierarchical “bottom-up approach” technique where each data point composed one cluster at the beginning and then paired with clusters that merged as they moved up in the hierarchy depending on their distance metrics was employed (Maimon and Rokach 2005). The distance measure was chosen “Single Linkage,” which meant the closest pair of elements that belonged to different clusters was taken into consideration while merging clusters. Moreover, “cosine similarity” was used as a similarity metric. However, obtained results included quite unbalanced two top clusters with 4 and 996 observations.

Next, k-means clustering, where k = 2, was adopted. Again “cosine similarity” metric was applied as the distance metrics among vectors. More balanced clusters, cluster_0 with 564 items and cluster_1 with 436 items, were achieved (Table 8). Subsequently, the five regular attributes, namely containsHashtag, containsLink, numberOfFollowers, numberOfCapitalLetters, containsPoliticWords and one

Table 8 Confusion matrix on predicting clusters

	True Cluster_1	True Cluster_0	Class Precision
Predicted Cluster_1	160	105	60.38%
Predicted Cluster_0	276	459	62.45%
Predicted total	436	564	
Class recall	36.70%	81.38%	
Accuracy	61.90 ± 3.96% (micro: 61.90%)		

special attribute, which is cluster_0 or cluster_1, were analyzed with the K-NN (k = 50) model. Results indicated the overall accuracy of the cross-validation as 61.9% (Table 8).

Sampling among the top highest re-Tweeted Tweets was believed to enclose similar major characteristics, hindering the aimed prediction capability of the model built. Therefore, Tweets with low re-Tweet incidents were included in further analysis to understand the content-based characteristic differences of highly re-Tweeted Tweets. Consequently, the training dataset was established to contain the top 1,000 highest re-Tweeted Tweets, where the number of re-Tweets for each Tweet ranged from 60 to 9,973. In addition, 1,000 random Tweets with each Tweet less than or equal to 5 re-Tweets were included in the dataset. Hence, 1,000 highly (if ≥ 60 re-Tweets, then label high) and 1,000 lowly (if ≤ 5 re-Tweets, then label low) re-Tweeted Tweets were included in the dataset with equal size bins. Subsequently, the five regular attributes, namely containsHashtag, containsLink, numberOfFollowers, numberOfCapitalLetters, and containsPoliticWords, were used in the learning model by using K-NN (k = 20). When tested with the cross-validation method, the built model’s overall accuracy is reported to be 83.80%, while both precision and recall values for each class were satisfactory (Table 9).

To further improve the dataset with regards to high and low classes, a follow-up model was built with the top 576 highest re-Tweeted Tweets (# of re-Tweets ≥ 105) and 548 of the least re-Tweeted Tweets (# of re-Tweets ≤ 2). In this way, the results were compared when the classes became more polarized among each other. K-NN (k = 20) classifier was executed with this dataset, and the overall accuracy was improved to 85.76% (Table 10).

As a result, an increase in the number of re-Tweets difference between the high and low classes in the dataset was shown to improve the accuracy of distinguishing classes

Table 9 Confusion matrix on predicting low and high labels with 2000 Tweets

	True high	True low	Class precision
Predicted high	863	187	82.19%
Predicted low	137	813	85.58%
Predicted total	1,000	1,000	
Class recall	86.30%	81.30%	
Accuracy	83.80 ± 2.14% (micro: 83.80%)		

Table 10 Confusion matrix on predicting low and high labels for more polarized classes

	True high	True low	Class precision
Predicted high	495	79	86.24%
Predicted low	81	469	85.27%
Predicted total	576	548	
Class recall	85.94%	85.58%	
Accuracy	85.76 ± 3.95% (micro: 85.77%)		

by the five attributes used (containsHashtag, containsLink, numberOfFollowers, numberOfCapitalLetters, and containsPoliticWords).

Identifying Class-Specific Terms. In information retrieval, the use of specific tokens (terms) leads to richer information on the class, particularly in the analysis of text data. Accordingly, topics (classes) were typically identified by finding the special words that characterize the relevant class following the approach proposed by Rajaraman and Ullman (2011). The importance of a word for a particular class was calculated based on its occurrence in a given class together with but less frequent occurrence in any other class. The determined class-specific terms for high and low classes were then used as another feature in testing an improved prediction capability. Accordingly, two new features were identified as:

- containsHighRelatedTerms: a Boolean value true or false
- containsLowRelatedTerms: a Boolean value true or false

These two new features were added to the training set that had 2,000 samples (1,000 low–1,000 high, the model’s data presented in Table 9). To summarize, the seven features used in this analysis involved containsHashtag, containsLink, containsHighRelatedTerms, containsLowRelatedTerms, numberOfCapitalLetters numberOfFollowers, and containsPoliticWords.

Use of K-NN ($k = 20$) classifier showed that accuracy did not improve much but remained at the same level of 83.80% (Table 9). When the new two additional features were added to the training set with 548 low and 576 high classes, accuracy increased from 85.76% (Table 10) to 86.12% (Table 11).

In order to capture the moderate number of re-Tweets, that is, $5 < \text{number of re-Tweets} < 60$, another 1,000 randomly selected representatively re-Tweeted Tweets were added to the training set. They were labeled as the low class since the focused prediction is on the top re-Tweet occurrences. As a result, the training set included a total of 3,000 Tweets, a combination of 2,000 low and 1,000 high Tweets. This was also reflective of the real Tweet eco-system since there were always more low levels of re-Tweeting than high. The confusion matrix on this dataset with K-NN ($k = 20$) is presented in Table 12. The overall accuracy of 75.10% on this dataset provided sound prediction accuracy in predicting whether a Tweet will get a high number of re-Tweets.

Some characteristic Tweet features were observed for a Tweet to lead towards extensive re-Tweeting. Accordingly, it is concluded that Tweets containing political

Table 11 Confusion matrix on predicting low and high labels with occurrence features

	True high	True low	Class precision
Predicted high	488	68	87.77%
Predicted low	88	480	84.51%
Predicted total	576	548	
Class recall	84.72%	87.59%	
Accuracy	86.12 ± 2.57% (micro: 86.12%)		

Table 12 Confusion matrix on predicting low and high labels with 3000 Tweets

	True high	True low	Class precision
Predicted high	612	359	63.03%
Predicted low	388	1 641	80.88%
Predicted total	1 000	2 000	
Class recall	61.20%	82.05%	
Accuracy	75.10 ± 2.38% (micro: 75.10%)		

words were more likely to amass high re-Tweeting when compared to Tweets without political words. The findings indicated that 32% of the highly re-Tweeted Tweets contained political words, while only 27% of the lesser re-Tweet receiving Tweets contained political words. Therefore, determining whether a Tweet contains a political token (word) was identified as a distinctive feature in predicting the likelihood for high re-Tweeting. Note that 1,000 Tweets, which had more than 60 re-Tweets, were labeled as high (more specifically, the number of re-Tweets for each Tweet we labeled high ranged from 60 to 9,973). Additionally, Tweets containing a hyperlink tend to be highly re-Tweeted by Turkish Twitter users. Moreover, containing class-specific terms (containsHighRelatedTerms and containsLowRelatedTerms) was another feature in predicting the likelihood for high re-Tweeting.

5 Discussion of Major Findings

As Tweet volume was an indicator of social trends taking place in real life, some common categories emerging in the dataset were identified. In order of importance, these categories were armed conflicts, religious and political sensitivities, and cultural traits such as in-group collectivism and paternalism.

Tweets containing Syria were not only posted by a small specific group of people but indeed attracted new individuals to post, with cumulative numbers of 175,000 users. The topic of Syria was reflected upon by a wide range of individuals as illustrated in the patterns of distinct daily users and Tweet numbers distributions.

Among the highly re-Tweeted top 50 Tweets (Table 2), Tweet1 received the highest number of re-Tweets, Tweet3 had the highest number of followers, Tweet6 and Tweet38 achieved the sharpest and smoothest slope, respectively (see Fig. 4).

In order to predict features that lead to high re-Tweet, a model was built by first with 500 top re-Tweeted Tweets, then with a total of 3,000 Tweets where a combination of 2,000 low and 1,000 high Tweets was used. This was more reflective of the real Tweet eco-system, which often included more low levels of re-Tweeting than the high. Classification of high and low Tweets was based on the number of re-Tweets received, where high was considered as having more than 60 re-Tweets. By using seven features, namely containsHighRelatedTerms, containsLowRelatedTerms, containsHashtag, containsLink, numberOfFollowers, numberOfCapitalLetters, and

contains `PoliticWords`; use of K-NN ($k = 20$) classifier showed that the best accuracy was achieved at the level of 86.12% (Table 11). It is often assumed that the `numberOfFollowers` would be a dominant regular attribute directly related to the number of re-Tweets. Upon closer examination, the analysis revealed that 800 of 2,000 low labeled Tweets—that is, less re-Tweeted Tweet posting accounts—had more than 10,000 followers. However, the developed model still achieved good prediction results. In other words, the developed learning model indicated that there was no direct relation between the number of followers of a Twitter account and the re-Tweets that the Tweets posted from this account would have. This finding also contributes to the research field in terms of correcting a common misassumption.

The mayors of Turkey did not use Twitter for transparent, participatory, and citizen-oriented public service delivery but rather used Twitter for self-promotion and political marketing (Sobaci and Karkin 2013). The results of the analysis conducted in this chapter illustrated armed fighting, religious and political sensitivities were identified as topics leading to high Tweet volumes. Accordingly use of Twitter by politicians to promote measures of public safety, religious and political stability messages could be a preferred approach for policymakers.

A government account that aims to inform the public on security issues would be interested in achieving the highest number of re-Tweets in the shortest time to be able to spread information as fast as possible to the widest possible public. It is expected that having a high number of followers would be sufficient based on common understanding. However, the conducted analysis in this chapter indicated that a combination of seven identified features needs to be considered in this regard.

6 Conclusions and Future Research

This chapter presents an analysis of Tweets in exploring the reflections on the Syria conflict among the Turkish public as frontier research in this field. Unlike most social conflict-based Twitter research that rather preferred to focus on digital activists by studying mostly a narrow group of people engaging online for social change, the chapter contributed to the literature by providing a wider perspective by including all available social media reflections from the public. The chapter has not only chosen to have a deep research focus for a specific group that has anti-war, anti-immigration, or anti-anything but included all posted Tweets over a year-long period from any perspective for ensuring a wider focus representative of the public's reflection. The chapter contributed to the theory by uncovering themes from a wide span of data that does not rely on any sampling but collection of all posted Tweets that included the word 'Suriye' (Syria in Turkish) and were publicly available. As part of the chapter's findings, it is shown that social media, which was launched as part of a leisure activity alternative, took a different role in serving as a medium for discussing societal concerns that can be as severe as war and immigration.

The analysis revealed that armed fighting, religious and political sensitivities within the Turkish public tend to inflate the volume of posted Tweets during the

Syrian conflict. In addition, Tweets containing Syria were not only posted by a small specific group of people but indeed attracted new individuals to post. The topic of Syria was reflected upon by a wider range of individuals.

A predictive model of 86.12% accuracy was built to classify high and low Tweets based on the number of re-Tweets received via seven features, namely numberOfFollowers, containsHighRelatedTerms, containsLowRelatedTerms, containsPoliticalWords, containsLink, numberOfCapitalLetters, and containsHashtag. As a surprising finding, the chapter concluded that there was no direct relation between the number of followers of a Twitter account and the re-Tweets that the Tweets posted from this account are likely to have. This finding provided an unexpected methodological contribution towards further research that utilizes Twitter and other social media data.

Policymakers can use Twitter to promote measures of public safety, religious and political stability messages effectively by considering the seven identified features in the model built. Policymakers should revisit the tendency to rely on traditional communication channels on public concern issues, and opportunities to utilize new mediums such as social media should be seized. Public concerns often involve citizens' voluntary contribution to gathering and spreading information, which Twitter offers widespread usage for (Gürbüz et al. 2017), especially in sound information seeking (Zha et al. 2015). For instance, in line with the propositions set forth by Tanes (2017) to use gameplay for improving earthquake precautions, other innovative novel communication approaches could be developed for efforts concerning issues surrounding the social conflict, such as integration of the immigrants via Twitter. Twitter is renowned for its impact in building social capital, especially for communication support (Son et al. 2016). E-marketers working on policy communication design can increase their reach, as it is reported in the case of the Danish Police's message diffusion study (Velde et al. 2015) by catering to citizens that engage in social media (Akar and Dalgic 2018) by utilizing corporate Twitter accounts (Lee and Kim 2018).

The retrieved data in this chapter's sampling by Twitter Stream API had some limitations since Twitter did not permit retrieval of all Tweets but instead allowed streaming of only a small fraction of the total volume of Tweets at any given moment³. Therefore the findings were only based on Tweets within this limitation. However, it is believed that the arrived results were applicable in understanding the Syrian conflict Tweets posted by the public in Turkey. Moreover, the majority of the Tweets sent daily had some rather low number of re-Tweets. Therefore, the whole dataset could not be used for the testing process as it would have led to an unbalanced distribution among the "low" and "high" classes, with a weaker learning model. Therefore, a subset for each class with the same size to test the model was selected. Another limitation that comes with Tweet data was related to the bot accounts on Twitter. A Twitter bot is an account organized to post automatic Tweets by using a special software program instead of a real Twitter user. In the analysis, the top 500 most re-Tweeted Tweets' users were manually checked to confirm that they were real

³ Retrieved from <https://dev.twitter.com/tags/streaming-api> on August 17th, 2016.

individuals, celebrities, media organizations, and state institutions that posted them. Hence the general limitation of bot accounts did not apply to the chapter's findings.

Acknowledgements This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. However, as non-monetary support, Sabancı University's servers were utilized for data collection.

References

- Akar, E., Dalgic, T.: Understanding online consumers' purchase intentions: a contribution from social network theory. *Behav. Inf. Technol.* **37**(5), 473–487 (2018)
- Alexa Homepage. <http://www.alexa.com/topsites/countries/TR>. Accessed 02 June 2020
- Aljazeera Homepage: Syria's civil war explained from the beginning. <https://www.aljazeera.com/news/2016/05/syria-civil-war-explained-160505084119966.html>. Accessed 07 July 2021
- Aras, N.E.G., Mencutek, Z.S.: The international migration and foreign policy nexus: the case of Syrian refugee crisis and Turkey. *Migr. Lett.* **12**(3), 193–208 (2015)
- Arcan, Z.: Human resource management in Turkey: current issues and future challenges. *Int. J. Manpow.* **22**(3), 252–260 (2001)
- Berger, J., Milkman, K.L.: What makes online content viral? *J. Mark. Res.* **49**(2), 192–205 (2012)
- Bircan, T., Sunata, U.: Educational assessment of Syrian refugees in Turkey. *Migr. Lett.* **12**(3), 226–237 (2015)
- Capece, G., Costa, R.: The new neighbourhood in the internet era: network communities serving local communities. *Behav. Inf. Technol.* **32**(5), 438–448 (2013)
- Çevik, M., Ozturkcan, S., Kasap, N.: Social media analytics: big data approach for Twitter. In: Paper Presented at the 20th Marketing Congress, Eskisehir, Turkey (2015)
- Chatty, D.: Research in Brief: The Syrian humanitarian disaster: disparities in perceptions, aspirations and behaviour in Lebanon, Jordan and Turkey (RSC Research in Brief 3), 04 December 2015
- Coletto, M., et al.: Sentiment-enhanced multidimensional analysis of online social networks: perception of the Mediterranean refugees' crisis. In: Paper Presented at the Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Davis, California (2016)
- Connor, P.: Most Displaced Syrians are in the Middle East, and About a Million are in Europe, 29 January 2018. <http://www.pewresearch.org/fact-tank/2018/01/29/where-displaced-syrians-have-resettled/>. Accessed 17 Feb 2021
- Dekker, R., Engbersen, G.: How social media transform migrant networks and facilitate migration. *Global Netw.* **14**(4), 401–418 (2014)
- Della Porta, D.: Comment on organizing in the crowd. *Inf. Commun. Soc.* **17**(2), 269–271 (2014)
- Fan, W., Gordon, M.D.: The power of social media analytics. *Commun. ACM* **57**(6), 74–81 (2014)
- Gabiam, N.: Humanitarianism, development, and security in the 21st century: lessons from the Syrian refugee crisis. *Int. J. Middle East Stud.* **48**, 382–386 (2016)
- García-Martín, J., García-Sánchez, J.-N.: Use of Facebook, Tuenti, Twitter and Myspace among young Spanish people. *Behav. Inf. Technol.* **34**(7), 685–703 (2015)
- Gerbaudo, P.: *Tweets and the Streets: Social Media and Contemporary Activism*. PB-Pluto Press, London (2012)
- GLOBE: Global Leadership and Organizational Behavior Effectiveness (2016). <http://globe.bus.sfu.ca/results/countries/TUR?menu=list>. Accessed 12 May 2021
- Gürbüz, H.G.A., Demir, T., Özcan, B.G., Kadak, M.T., Poyraz, B.C.: Use of social network sites among depressed adolescents. *Behav. Inf. Technol.* **36**(5), 517–523 (2017)

- Heisbourg, F.: The strategic implications of the Syrian refugee crisis. *Survival* **57**(6), 7–20 (2015)
- Hofstede, G.: *Culture's Consequences: International Differences in Work-Related Values*. Sage, Thousand Oaks (1980)
- House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W., Gupta, V.: *Culture, Leadership and Organizations: The GLOBE Study of 62 Societies*. Sage Publications Inc., Thousand Oaks (2004)
- Howard, B.: Analyzing online social networks. *Commun. ACM* **51**(11), 14–16 (2008)
- Issa, P.: Syrian civil war: Five ways the conflict has changed the world. Independent (2016). <http://www.independent.co.uk/news/world/middle-east/syrian-civil-war-isis-how-it-changed-the-world-refugee-crisis-a6928796.html>. Accessed 03 Sept 2021
- Jacobs, A.: The pathologies of big data. *Commun. ACM* **52**(8), 36–44 (2009)
- Jansen, F.: Digital activism in the Middle East: mapping issue networks in Egypt, Iran, Syria and Tunisia. *Knowl. Manag. Dev. J.* **6**(1), 37–52 (2010)
- Kabasakal, H., Bodur, M.: *Leadership, values and institutions: The case of Turkey*. Bogazici University Research Papers, Istanbul, Turkey (1998)
- Karasapan, O.M.: The impact of Syrian businesses in Turkey (2016). <https://www.brookings.edu/2016/03/16/the-impact-of-syrian-businesses-in-turkey>. Accessed 14 Mar 2021
- Kourouthanassis, P., Lekakos, G., Gerakis, V.: Should I stay or should I go? The moderating effect of self-image congruity and trust on social networking continued use. *Behav. Inf. Technol.* **34**(2), 190–203 (2015)
- Lee, A.R., Kim, K.K.: Customer benefits and value co-creation activities in corporate social networking services. *Behav. Inf. Technol.* **37**(7), 675–692 (2018)
- Leuven, S.V., Heinrich, A., Deprez, A.: Foreign reporting and sourcing practices in the network sphere: a quantitative content analysis of the Arab Spring in Belgian news media. *New Media Soc.* **17**(4), 559–573 (2015)
- Liu, X.Y., Wu, J., Zhou, Z.H.: Exploratory undersampling for class-imbalance learning. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **39**(2), 539–550 (2008)
- Maimon, O., Rokach, L.: *Data Mining and Knowledge Discovery Handbook*. Springer, New York (2005). <https://doi.org/10.1007/978-0-387-09823-4>
- Maitland, C., Xu, Y.: A social informatics analysis of refugee mobile phone use: a case study of Za'atari Syrian refugee camp. In: Paper Presented at the TPRC 43: The 43rd Research Conference on Communication, Information and Internet Policy Paper, 31 March 2015
- MEE: Economic effect of Syrian war at \$35bn: World Bank (2016). <http://www.middleeasteye.net/news/economic-effect-syrian-war-35bn-world-bank-1953497162>. Accessed 18 June 2021
- Metallo, C., Agrifoglio, R.: The effects of generational differences on use continuance of Twitter: an investigation of digital natives and digital immigrants. *Behav. Inf. Technol.* **34**(9), 869–881 (2015)
- Molaei, H.: Discursive opportunity structure and the contribution of social media to the success of social movements in Indonesia. *Inf. Commun. Soc.* **18**(1), 94–108 (2015)
- Ogan, C., Varol, O.: What is gained and what is left to be done when content analysis is added to network analysis in the study of a social movement: twitter use during Gezi Park. *Inf. Commun. Soc.* **20**(8), 1220–1238 (2017)
- Orhan, O.: Effects of the Syrian refugees on Turkey, vol. 195 (2015). http://tese.org.tr/wp-content/uploads/2015/11/Effects_Of_The_Syrian_Refugees_On_Turkey.pdf
- Ozturkcan, S., Kasap, N., Çevik, M., Zaman, T.: An analysis of the Gezi park social movement tweets. *Aslib J. Inf. Manag.* **69**(4), 426–440 (2017)
- Ozturkcan, S., Kasap, N., Tanaltay, A., Ozdinc, M.: Analysis of tweets about football: 2013 and 2018 leagues in Turkey. *Behav. Inf. Technol.* **38**(9), 887–899 (2019)
- Penney, J., Dadas, C.: (Re)Tweeting in the service of protest: digital composition and circulation in the occupy wall street movement. *New Media Soc.* **16**, 74–90 (2014)
- Peters, K., Chen, Y., Kaplan, A.M., Ognibeni, B., Pauwels, K.: Social media metrics—a framework and guidelines for managing social media. *J. Interact. Mark.* **27**(4), 281–298 (2013)
- Phillips, C.: The impact of Syrian refugees on Turkey and Jordan. *World Today* **68**(8), 34–37 (2012). <https://www.chathamhouse.org/publications/twt/archive/view/186289>. Accessed 15 Apr 2021

- Rajaraman, A., Ullman, J.D.: *Mining of Massive Datasets*. Cambridge University Press, Cambridge (2011)
- Richter, F.: Retrieved from <https://www.statista.com/chart/1642/regional-breakdown-of-twitter-users/>. Accessed 12 Aug 2021
- Rui, H., Liu, Y., Whinston, A.: Whose and what chatter matters? The effect of tweets on movie sales. *Decis. Support Syst.* **55**(4), 863–870 (2013)
- Sakaki, T., Okazaki, M., Matsuo, Y.: Earthquake shakes Twitter users: real-time event detection by social sensors. In: Paper Presented at the Proceedings of the 19th International Conference on World Wide Web, Raleigh, North Carolina, USA (2010)
- Sakaki, T., Okazaki, M., Matsuo, Y.: Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Trans. Knowl. Data Eng.* **25**(4), 919–931 (2013)
- Sandoval-Almazan, R., Gil-Garcia, J.R.: Towards cyberactivism 2.0? Understanding the use of social media and other information technologies for political activism and social movements. *Gov. Inf. Q.* **31**, 365–378 (2014)
- Savage, N.: Twitter as medium and message. *Commun. ACM* **54**(3), 18–20 (2011)
- Seo, H.: Visual propaganda in the age of social media: an empirical analysis of twitter images during the 2012 Israeli-Hamas conflict. *Vis. Commun. Q.* **21**(3), 150–161 (2014)
- Shi, Z., Rui, H., Whinston, A.B.: Content sharing in a social broadcasting environment: evidence from Twitter. *MIS Q.* **38**(1), 123 (2014)
- Singhal, A.: Modern information retrieval: a brief overview. *Bull. IEEE Comput. Soc. Tech. Committee Data Eng.* **24**, 35–43 (2001)
- Smith, K.: *Marketing: 96 Amazing Social Media Statistics and Facts for 2016* (2016). <https://www.brandwatch.com/2016/03/96-amazing-social-media-statistics-and-facts-for-2016/>. Accessed 12 Sept 2021
- Sobaci, M.Z., Karkin, N.: The use of Twitter by mayors in Turkey: tweets for better public services? *Gov. Inf. Q.* **30**(4), 417–425 (2013)
- Son, J.-E., Lee, S.-H., Cho, E.-Y., Kim, H.-W.: Examining online citizenship behaviours in social network sites: a social capital perspective. *Behav. Inf. Technol.* **35**(9), 730–747 (2016)
- Starbird, K., Palen, L.: (How) will the revolution be re-tweeted?: information diffusion and the 2011 Egyptian uprising. In: Paper Presented at the Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, Seattle, Washington, USA (2012)
- Stieglitz, S., Dang-Xuan, L.: Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *J. Manag. Inf. Syst.* **29**(4), 217–248 (2013)
- Subramani, M.R., Rajagopalan, B.: Knowledge-sharing and influence in online social networks via viral marketing. *Commun. ACM* **46**(12), 300–307 (2003)
- Sunpuwan, M., Niyomsilpa, S.: Perception and misperception: Thai public opinions on refugees and migrants from Myanmar. *J. Popul. Soc. Stud.* **21**(1), 47–58 (2012). <http://repository.li.mahidol.ac.th/dspace/bitstream/123456789/2927/2/pr-ar-malee-2012.pdf>
- Tanes, Z.: Shall we play again? The effects of repetitive gameplay and self-efficacy on behavioural intentions to take earthquake precautions. *Behav. Inf. Technol.* **36**(10), 1037–1045 (2017)
- Theocharis, Y., Lowe, W., Deth, J.W., van García-Albacete, G.: Using Twitter to mobilize protest action: online mobilization patterns and action repertoires in the occupy wall street, Indignados, and Aganaktismenoi movements. *Inf. Commun. Soc.* **18**(2), 202–220 (2014)
- Tremayne, M.: Anatomy of protest in the digital era: a network analysis of twitter and occupy wall street. *Soc. Mov. Stud.* **13**(1), 110–126 (2013)
- Tufekci, Z.: Not this one: social movements, the attention economy, and microcelebrity networked activism. *Am. Behav. Sci.* **5**(57), 848–870 (2013)
- Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welp, I.M.: Predicting elections with Twitter: what 140 characters reveal about political sentiment. In: Paper Presented at the Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (2010)
- Tyshchuk, Y., Wallace, W.A., Li, H., Ji, H., Kase, S. E.: The nature of communications and emerging communities on twitter following the 2013 Syria Sarin gas attacks. In: Paper presented at the 2014 IEEE Joint Intelligence and Security Informatics Conference, 24–26 September 2014

- UNHCR (2016). <http://data.unhcr.org/syrianrefugees/country.php?id=224>
- UNHCR Homepage (2019). <https://emergency.unhcr.org/entry/55772/refugee-definition>. Accessed 23 Sept 2021
- Velde, B., Meijer, A., Homburg, V.: Police message diffusion on Twitter: analysing the reach of social media communications. *Behav. Inf. Technol.* **34**(1), 4–16 (2015)
- Wall, M., Otis Campbell, M., Janbek, D.: Syrian refugees and information precarity. *New Media Soc.* **19**(2), 240–254 (2015)
- Watters, P.A., Ziegler, J.: Controlling information behaviour: the case for access control. *Behav. Inf. Technol.* **35**(4), 268–276 (2016)
- Worldometers Homepage. <http://www.worldometers.info/world-population/turkey-population/>. Accessed 21 Aug 2021
- Wyer, R.S., Chiu, C., Hong, Y.: *Understanding Culture: Theory, Research, and Application*. 1st edn. Psychology Press, New York (2009)
- Zha, X., Zhang, J., Yan, Y., Zha, D.: Sound information seeking in Web 2.0 virtual communities: the moderating effect of mindfulness. *Behav. Inf. Technol.* **34**(9), 920–935 (2015)