Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 752 – 764
© TÜBİTAK
doi:10.3906/elk-1804-9

# Refugees' social media activities in Turkey: a computational analysis and demonstration method

**Abdullah BÜLBÜL**[1*], **Salah HAJ ISMAIL**[2]
[1]Computer Engineering, Faculty of Engineering and Natural Sciences, Ankara Yıldırım Beyazıt University, Ankara, Turkey
[2]Civil Engineering, Faculty of Engineering and Natural Sciences, Ankara Yıldırım Beyazıt University, Ankara, Turkey

**Received:** 02.04.2018 • **Accepted/Published Online:** 01.11.2018 • **Final Version:** 22.03.2019

**Abstract:** This study performs a data analysis on refugees in Turkey based on their social media activities. In order to achieve this, we first propose a method to find their relevant public accounts and collect their activities generating a dataset. Then, we perform spatial and temporal analysis over this dataset to shed light on the most important topics and events discussed in social networks. We present the results graphically for ease of understanding and comparison. Our results indicate that we can reveal the most shared topics over a specific time and place as well as the change of pattern in refugees' activities through their reflection on social media. Moreover, this dataset facilitates a number of further and deeper analyses of the refugees in Turkey.

**Key words:** Information retrieval, social media, refugees

## 1. Introduction

Eight years ago, the wave of Arab spring crossed many Arab countries. Social media usage marked this period not only as an accelerator of the Arab spring but also as a means of interaction for people to respond to the events they experienced or faced [1]. Later, the events in these countries forced millions of people to leave their homes and become refugees. For instance, the number of Syrian refugees surpassed 5 million in the world according to UNHCR. Turkey currently accommodates more than 3.5 million of these refugees [2]. A dataset which includes the social media activities of refugees in Turkey has the potential to facilitate a number of future studies. Recent application areas that can benefit from such a dataset include complex social analysis [3, 4], knowledge communication [5], and link prediction [6, 7]. Thus, an indicative dataset could be established through data retrieval from social networks, which could help to present the events or topics that occupied the refugees in their refuge countries, where access through traditional methods such as questionnaires or interviews might be difficult.

This paper presents a tool and methodology to collect data from Twitter accounts, filter them to extract the refugees' accounts, then trace their tweets back to collect a dataset to be analyzed against time and location across Turkey, defining the trending topics related to the issues of their home country and new refuge country. In results and discussion, those trends are shown graphically which eases better understanding and comparison of multidimensional analysis.

The rest of the paper presents related works, explains our method of collecting data from the social media,

*Correspondence: abulbul@ybu.edu.tr

and the details of these collected data. We then analyze the nature of these data and potential difficulties that can be faced when retrieving and statistically examining them. Finally, we present several preliminary outcomes and discuss more detailed future potential fields to analyze.

## 2. Related works

Being able to monitor others' lives and notions regarding daily events or opinions concerning political issues, products, or even meals in the simplest terms is made possible via microblogging sites such as Facebook, Twitter, and Tumbler in which enormous quantity of data is continuously released by the social sensors or users. In the literature, there are numerous studies exploiting social media data for various purposes: to reveal the effect of a political or social event on society, to receive feedback concerning a product from customers of a company, to determine the location of an accident or disaster for emergency assistance, and even for early detection and analysis of epidemics. Since the aim in this study was to understand the problems and the needs of refugees based on the collected data from their tweets, we here present related works in which a long-term phenomenon, that may or may not be directly related to refugees, was analyzed using data from social media.

Elections might be considered one of the most popular worldwide long-term events, where candidates propagate their ideas, promises, and projects during the campaigns. During these times, social media produces vast comments and news concerning the candidates' statements; which may be categorized as positive, negative, or objective depending on the sentiment it comprises. Quite a few studies exist in the literature aiming to predict the elections, in terms of forecasting the winner of seats in the parliament and to understand how tweets reflect political sentiment in society [8–11]. Other studies question the reliability of this predictive research due to the popularity of social media-based analysis and the easy nature of the proposed methods. For instance, Gayo-Avello [12] claimed that the predicted results obtained via Twitter data had been exaggerated.

The effect of social media on social bursts is also undeniable. Government acts, such as disconnecting online sources to prevent organization and communication of opponent groups can be strong evidence on the power of microblogging sites where people can quickly become organized to respond to a social event. Because of the conflict in Syria, many citizens have migrated abroad since 2011. The significant population of these refugees around the world gave rise to several academic studies addressing the issue from various points of view. By means of collected tweets including #refugeesnotwelcome hashtag, Kreis [13] explored the ideas of users towards refugees. Another study [14] also exploited tweets including the same hashtag to understand the portrayal of male Syrian refugees on social media. The future attitudes of users based on past posts are also attempted to be predicted from social media data. As an example, Magdy et al. [15] attempted to predict the attitudes of US Twitter users towards Islam and Muslims subsequent to the tragic Paris terrorist attacks that occurred on November 13, 2015. In an interesting research study, Darwish et al. [16] showed the propagation power of Twitter on users via ”seminar users” who are social media users engaged in propaganda in support of a political entity.

## 3. Method used

Using Twitter API[1], we analyzed public Twitter activity to retrieve information from refugee-related social media accounts. For that purpose, the process was divided into four steps: Firstly, we developed a method to determine the accounts of Syrian refugees in Turkey. Secondly, we traced back the chosen users' accounts and

[1] https://developer.twitter.com/en/docs

collected all the possible tweets that were allowed under Twitter conditions. Thirdly, we statistically analyzed the data retrieved and classified them into groups by location and year. Finally, a trend analysis was conducted to understand the most important issues the refugees discussed over time and place. All the results were presented graphically and discussed to better understand the technical limitations faced with this process. In addition, the social aspects we retrieved suggest deeper and multidisciplinary future studies.

### 3.1. Determining the accounts

In Turkey, the majority of citizens communicate in Turkish. Therefore, the first heuristic we can use to determine the refugee-owned accounts is to check their languages. In Twitter, users are associated with a language. Of the registered foreigners in Turkey, 90% are refugees, the majority of refugees (95% [17]) have Arabic as their mother tongue, and of the remaining 10% of foreigners, 80% also speak Arabic, so 91.44% [2] of the Twitter accounts in Turkey whose main language is Arabic are expected to belong to refugees.

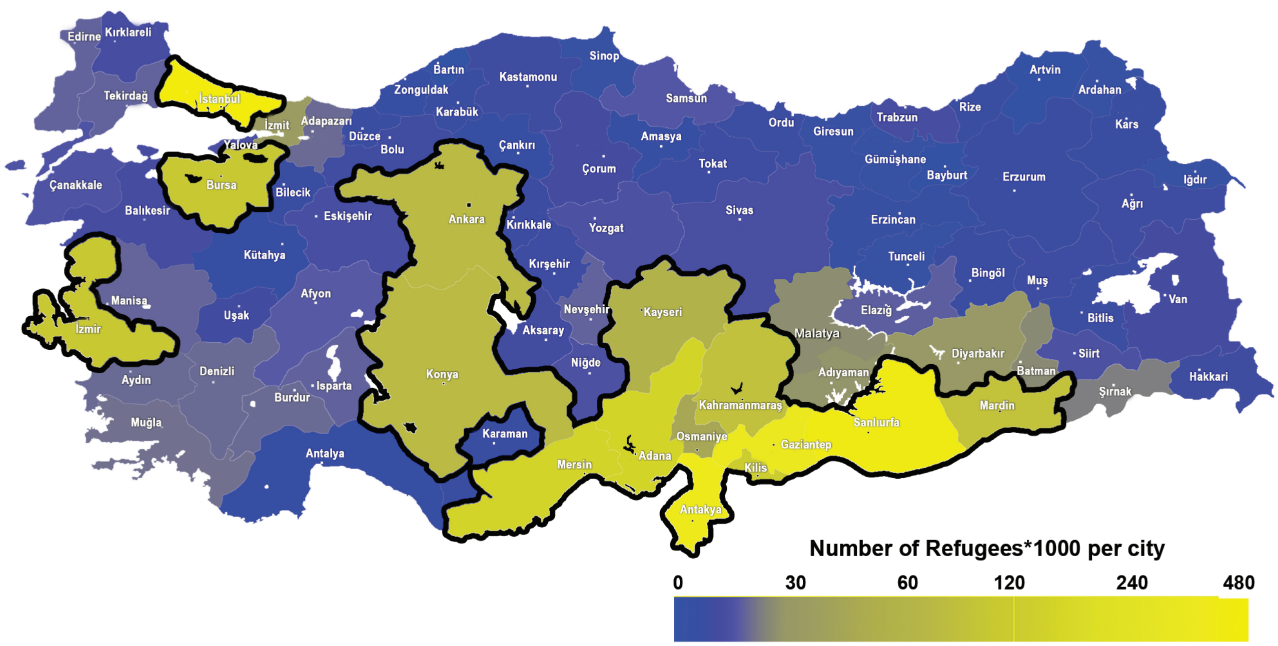


**Figure 1**. Distribution of Syrian refugees in Turkish cities generated according to the statistics in [18] using a logarithmic color code to effectively visualize the number of refugees in cities.

Syrians make up the majority of refugees in Turkey; therefore, we checked the number of the registered Syrian refugees in each city. Figure 1 shows their distribution in Turkey. The image was generated according to the statistics in [18] using a logarithmic color code to effectively visualize the number of refugees in cities.

From Figure 1, it is clear that the majority of the refugees are hosted in specific cities. The bold border shows the 15 cities chosen in our study and these cities accommodate 90% of the Syrian refugees while they only have 48% of the total Turkish population. Table 1 shows the number of refugees along with the populations of the chosen cities and the ratio of refugees to the city populations.

Twitter API allows searching for recent (seven days) tweets according to key words, location, and the language used. To avoid a biased collection of data, we did not provide any key words. We searched for Arabic

[2] $91.44 = 0.90 \times 0.95/(0.90 \times 0.95) + (0.10 \times 0.80)$

**Table 1**. Main cities hosting refugees.

| | Adana | Ankara | Bursa | Gaziantep | Hatay | İstanbul | İzmir | Kahramanmaraş |
|---|---|---|---|---|---|---|---|---|
| No. of refugees | 150,790 | 73,198 | 106,000 | 329,670 | 384,024 | 479,555 | 108,306 | 90,100 |
| Population | 2,201,670 | 5,346,518 | 2,901,396 | 1,974,244 | 1,555,165 | 14,804,116 | 4,223,545 | 1,112,634 |
| Ratio | 6.85% | 1.37% | 3.68% | 16.70% | 24.69% | 3.24% | 2.58% | 8.11% |
| | Kayseri | Kilis | Konya | Mardin | Mersin | Osmaniye | Şanlıurfa | Total |
| No. of refugees | 59,938 | 124,000 | 73,445 | 94,340 | 146,931 | 43,773 | 420,532 | 2,685,669 |
| Population | 1,358,980 | 130,825 | 2,161,303 | 796,237 | 1,773,852 | 522,175 | 194,627 | 42,803,287 |
| Ratio | 4.34% | 95.15% | 3.40% | 11.85% | 8.28% | 8.38% | 21.67% | 6.28% |

tweets in specific locations shown in Figure 2. To determine these coordinates and radii given in the figure, we checked the main neighborhoods of Syrian refugees in each studied city. Note that, in Twitter, the location is preferentially determined by Geotagging API, but it falls back to what users specify as their location in their profiles (i.e. cities, districts, etc.). In our initial attempt of having compact radii in each city, we had cases in which coarse location assigned by users did not allow Twitter to define any city within the suggested radius. In such cases, we needed to increase the search radii in several cities. Users who did not specify their location in their profiles were not included in the analysis. Figure 2 shows the initial and final search radii under each location.

We collected the recent tweets in Arabic from those specified regions. For practical purposes, we limited the search to 1000 tweets per region. We kept these tweets in json format. As is known, the structure of a tweet includes information on time, Tweet ID, tweet or retweet, language, user location, and user ID. We then extracted the individual user IDs posting these tweets. As one account can post multiple tweets, we had a fewer amount of user accounts than the number of tweets. We performed this procedure twice with a 4-day interval to increase the number of users. As a result, we collected a total of 5707 Twitter users who had been active recently. Table 2 shows the number of users discovered in each region.

**Table 2**. Data collected: users and tweets per city.

| | Adana | Ankara | Bursa | Gaziantep | Hatay | İstanbul | İzmir | Kahramanmaraş |
|---|---|---|---|---|---|---|---|---|
| No. of users | 160 | 1050 | 2023 | 535 | 12 | 1183 | 78 | 35 |
| No. of tweets | 127,031 | 2,297,028 | 4,543,793 | 1,139,878 | 13,874 | 2,910,948 | 107,263 | 13,538 |
| Ratio Rt/T | 71.7% | 25.9% | 54.5% | 84.6% | 22.8% | 67.0% | 40.7% | 30.9% |
| | Kayseri | Kilis | Konya | Mardin | Mersin | Osmaniye | Şanlıurfa | Total |
| No. of users | 62 | 3 | 137 | 30 | 269 | 22 | 108 | 5707 |
| No. of tweets | 55,275 | 558 | 158,845 | 24,130 | 461,598 | 11,811 | 156,736 | 12,022,306 |
| Ratio Rt/T | 35.8% | 47.1% | 46.6% | 74.0% | 69.4% | 41.7% | 72.5% | 55.5% |

### 3.2. Tracing back and filtering out irrelevant accounts

Twitter API allows access to up to the last 3200 of a user's tweets including retweets. Therefore, we gathered tweets of each user we extracted in the previous step until the limit was reached or there were no more tweets from that user. In one query, it was possible to access only 200 tweets; thus, we ran multiple queries to collect the maximum possible number of tweets. Table 2 shows the number of tweets collected from each city. Among these accounts, there are some which do not belong to individual users but to press or companies for instance.

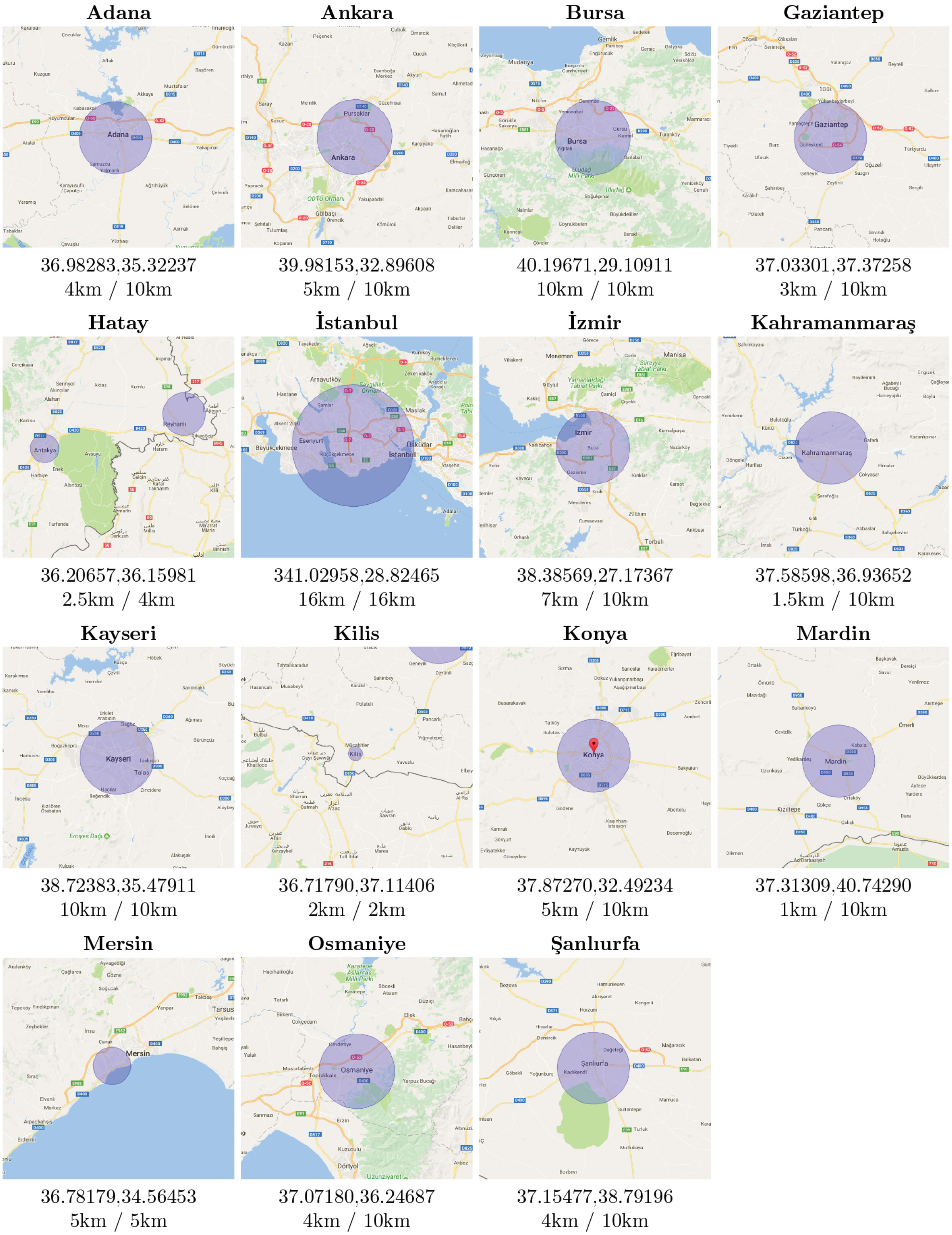


**Figure 2**. The zones studied. For each city, the circles depict the region for which Twitter activity is downloaded. Images are on the same scale.

These accounts mostly posted with a very high frequency also including a big ratio of retweets. This information is helpful to differentiate the users we are interested in to analyze in this study.

Figure 3 shows that the increase in Syrian refugees number in Turkey started from 2013. In order to analyze the development of refugee-related issues in post-"Arab Spring" years in Turkey, we focus more on the users for whom we had data that went back at least before 2014. Table 3 shows the number of accounts with the date of their oldest tweets collected.

**Table 3**. Number of users with their oldest known activity.

| Year | 2018 | 2017 | 2016 | 2015 | 2014 | 2013 | 2012 | 2011 | 2010 |
|---|---|---|---|---|---|---|---|---|---|
| No. of users | 876 | 2997 | 815 | 397 | 207 | 194 | 147 | 63 | 22 |

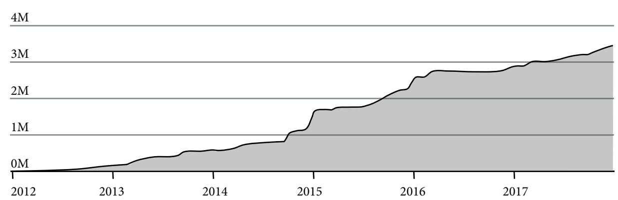


**Figure 3**. The total number of Syrian refugees in Turkey from 2012 until now [19].

### 3.3. Analysis

Firstly, we excluded the accounts which have all the tweets in 2018. Only tweets up to the end of 2017 were analyzed, since in 2018, millions of tweets (3,678,739) were retrieved in less than 20 days (See Table 4). This part of the data is somewhat unreliable due to the fact that tweeting 3200 tweets in those days means more than 160 tweets per day. This implies that these tweets neither coming from a real user nor expressing an individual user opinion.

In order to trace back and analyze the refugees' conditions and activities before and after coming to Turkey, we excluded the Twitter accounts created or have all the activities after 2014 since these accounts are incapable of reflecting the continuous change of refugees' conditions and needs.

With the aforementioned approach, out of 5707 accounts and 12 million tweets; the accounts and tweets to be analyzed decreased to 633 and almost 800 thousand, respectively. These tweets include 435,378 tweets and 336,753 RTs. In this study, we are concerned with the direct opinions of refugees. Thus, we do not include RTs in our analysis. We found out the frequency of occurrence of each word in the combined text from all tweets. It is possible to weigh a tweet according to its retweet and like count, which may give biased results towards popular accounts that we avoided in this study. Since most of the conjunctions, pronouns, etc. in Arabic have less than three letters, we discarded any word which has less than three letters to obtain useful and meaningful words.

Conducting an analysis based on the direct count of words misleadingly favors the commonly used words of daily life rather than the representative words that reflect a specific event or topic. To avoid highlighting such common words, we use the following simple approach: Assume $T$ is the set representing all the text combined (in our case it contains all tweets collected from 2012 to 2017) and $|T|$ is the total number of words in $T$. For each word, the base frequency is simply calculated as follows.

$$f_{base}(w) = |w_T|/|T|, \tag{1}$$

**Table 4**. Tweets per year and the ratio of retweets.

| | Adana | Ankara | Bursa | Gaziantep | Hatay | İstanbul | İzmir | Kahramanmaraş |
|---|---|---|---|---|---|---|---|---|
| 2018 | 63,250 (87%) | 804,872 (14%) | 770,299 (64%) | 707,998 (94%) | 553 (42%) | 1,029,079 (84%) | 24,389 (34%) | 1100 (62%) |
| 2017 | 48,196 (55%) | 1,279,007 (29%) | 2,800,561 (55%) | 355,777 (76%) | 5921 (10%) | 1,593,994 (60%) | 63,366 (46%) | 9767 (29%) |
| 2016 | 8593 (61%) | 117,600 (45%) | 520,959 (53%) | 35,843 (45%) | 3952 (36%) | 176,940 (43%) | 10,838 (40%) | 996 (32%) |
| 2015 | 5496 (60%) | 45,407 (49%) | 186,244 (47%) | 18,632 (29%) | 3171 (25%) | 59,848 (47%) | 5671 (24%) | 220 (31%) |
| 2014 | 1484 (59%) | 23,979 (67%) | 131,849 (38%) | 9878 (17%) | 113 (47%) | 29,421 (36%) | 2398 (30%) | 697 (30%) |
| 2013 | 12 (00%) | 19,379 (77%) | 90,014 (28%) | 7078 (18%) | 101 (64%) | 14,457 (27%) | 380 (07%) | 545 (07%) |
| 2012 | 0 ( -%) | 5859 (63%) | 40,144 (33%) | 3338 (26%) | 35 (11%) | 5768 (28%) | 141 (06%) | 186 (09%) |
| | Kayseri | Kilis | Konya | Mardin | Mersin | Osmaniye | Şanlıurfa | Total |
| 2018 | 8987 (37%) | 45 (44%) | 33,547 (64%) | 13,646 (85%) | 187,018 (83%) | 1609 (55%) | 32,347 (79%) | 3,678,739 (66%) |
| 2017 | 27,853 (39%) | 68 (12%) | 85,525 (46%) | 9072 (57%) | 214,958 (62%) | 8166 (28%) | 106,439 (72%) | 6,608,670 (52%) |
| 2016 | 9001 (52%) | 203 (01%) | 23283 (33%) | 231 (35%) | 30,391 (66%) | 1939 (85%) | 11,935 (57%) | 952,704 (49%) |
| 2015 | 4546 (09%) | 74 (86%) | 5866 (42%) | 196 (35%) | 11,475 (70%) | 86 (91%) | 5112 (83%) | 352,044 (47%) |
| 2014 | 1276 (16%) | 168 (100%) | 4466 (30%) | 131 (54%) | 6992 (26%) | 11 (27%) | 368 (04%) | 213,231 (39%) |
| 2013 | 1308 (08%) | 0 ( -%) | 4388 (30%) | 65 (54%) | 7552 (16%) | 0 ( -%) | 442 (05%) | 145,721 (33%) |
| 2012 | 1794 (11%) | 0 ( -%) | 1551 (04%) | 676 (99%) | 2822 (08%) | 0 ( -%) | 37 (00%) | 62,351 (33%) |

where $w$ is a word in $T$ and $|w_T|$ is the number of occurrences of $w$ in $T$. To obtain the representative words of a specific domain, e.g., a specific time or location, we collected the tweets in this domain creating a subset of $T$ named $D$. Similarly, we calculated the domain frequency of each word in $D$:

$$f_{domain}(w) = |w_D|/|D| \tag{2}$$

Then, the difference of domain-specific frequency of a word compared to its base frequency gave us the uniqueness of this word in the specific domain.

$$u(w_D) = f_{domain}(w) - f_{base}(w), \tag{3}$$

This approach highlights the representative words in a domain of interest and suppresses the misleadingly high frequency of common words, e.g., pronouns, conjunctions, and suffixes. An alternative approach to eliminate these words is using stemmers and stop words as in [19]. However, this approach creates the risk of eliminating some meaningful and highly important representative words. For instance, even though عبر (Abra), which means through or across in English, is a representative word for October 2016; it would be eliminated because it is a preposition.

In addition to that, using stemmers would increase the ambiguity of Arabic texts. Arabic is already a highly ambiguous language because of its structure using short vocals and movements to change the meaning of a word. With the usual tweeting mode in Arabic which does not use any of these short vocals or movements, an Arabic stemmer could misunderstand the different parts of speech and eliminate important words. For example, the same word (abra) will completely change into (Abara) not by adding letters but only by changing the tone, which changes the word into a verb with the meaning 'crossed'. Moreover, stemmers have difficulty in stemming words correctly in Arabic [20]; they remove letters considering them as suffixes or prefixes while they are originally part of the word. For example, معالم ("eatures" in English) stemmer will remove the 'Mim' at the beginning of the word and it will be عالم which could mean World or Scientist according to the tone dilemma mentioned before.

Because of the issues created by using Arabic stemmers, we preferred using the suggested procedure to choose the meaningful words by considering their uniqueness.

## 4. Results and discussion

The collected dataset was analyzed per various domains. Firstly, we made an analysis on a yearly basis throughout Turkey, between 2012 and 2017. The results are presented in Figure **??**, where the word clouds are created to show the trends using the uniqueness score explained previously while the English word clouds show manual translations of the Arabic ones.

The change in topics is notable across the years. Deep involvement in the Syrian crisis is apparent in the first four years, while issues related with obstacles in Turkey, such as work, Kızılay, dreaming of returning, beauty, can be noticed in later years. Achievements of Turkey, the failed coup attempt, and taking a decision of immigrating to Europe or staying in Turkey were the main topics in the last years.

Another important change noticed clearly is the change of language used in the tweets by the same accounts. Starting by using Turkish words in the Arabic alphabet, e.g., Para which means "money", then using Turkish language and alphabet in the tweets reflects that those users have started to learn Turkish. This change indicates that their integration process in the society has become deeper and stronger when their stay took longer than they expected at the beginning.

Secondly, in order to break it down, we changed our domain to be on monthly instead of yearly basis, and as an example, we analyzed the tweets in 2016 per month. The results are shown in Figure 5 with the English translation. The radii of monthly word clouds are determined by the size of Twitter activities in the respective months. Thus, we can understand that in July and August, the activities were doubled in comparison with the other months. This is mainly because of the special events in these two months. In July, the coup attempt in Turkey, which failed due to people's unity and determination, is clearly seen in the word cloud while in August, the great attack on Aleppo took place and occupied the refugees' minds and interaction in this month. The prevalence of the word Aleppo in 2016 was due to the high occurrence of Aleppo events in five months of this year.

Thirdly, in a location-based analysis, the clear differences in refugees' situation and trending issues among Turkish cities could be observed. Figure 6 shows the trending tweets in each chosen city in 2016 with their


**Figure 4**. Generated word clouds for years 2012 to 2017. Same word clouds for each year presented: original Arabic words (right), 20 most frequent words translated into English (left).

English translation. We can notice the difference in topics between a touristic city like İstanbul, where words like "coffee" and "beautiful" are the most occurring words, while cities like Gaziantep and Şanlıurfa, which are closer to the Syrian border, were more reflective of the events in Syria. Words like "Aleppo" and "Besieged" are the obvious ones.

In addition to that, it should be noticed that we excluded retweets from the analysis. A future comparative analysis between tweets and retweets could show other results or even different trends since we assume that the tweets are directly expressing individual views, while retweets could belong to press or seminar users [16] trying to orient the public to collective political opinions.

Different fields and topics can be examined using this dataset in order to retrieve more information about specific issues, such as services to refugees, legislative, and logistics of their stay in the refuge countries, and reactions to political changes and treatment in hosting countries. Moreover, other than the refugee-related studies, the same combination of techniques presented in this study is applicable to analysis in various fields. For instance, tracing the public opinion across years about specific topics, e.g., education, economy, and politics

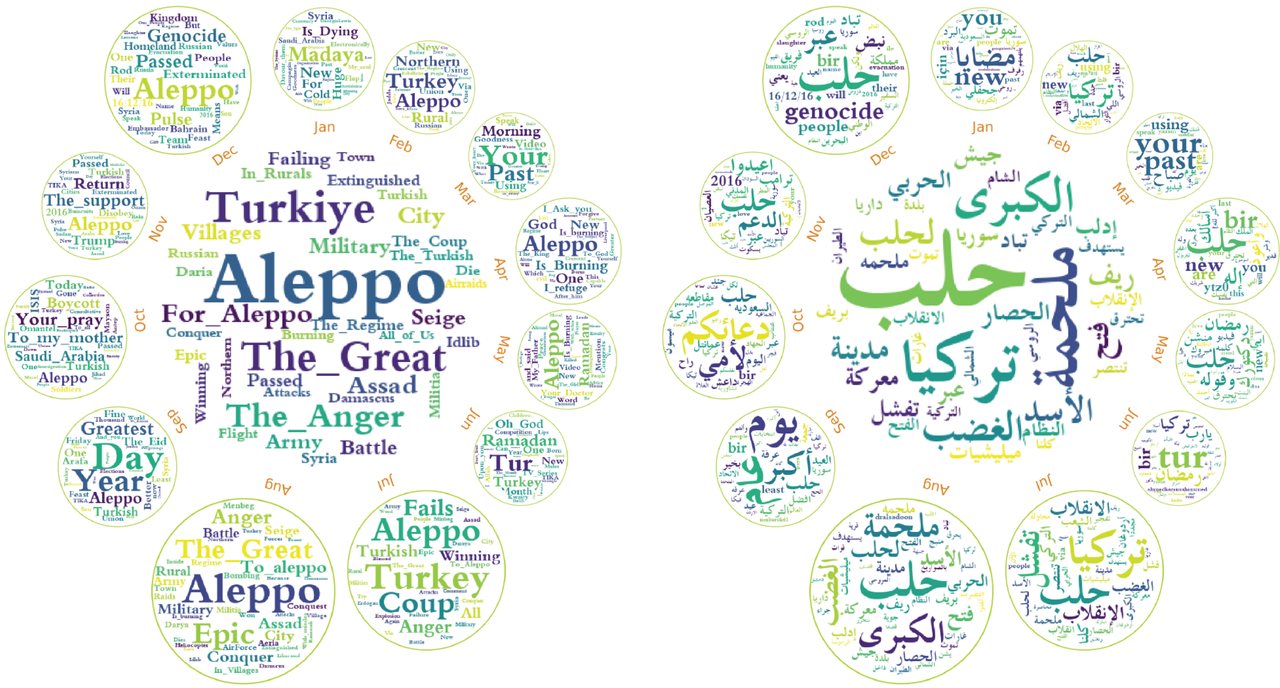


**Figure 5**. Word cloud showing the monthly analysis results in 2016: original (right) and translation into English (left). Although Aleppo is the notable topic throughout the year, each month has a different trend. For instance, Turkey and coup were the dominating topics in July.

would be beneficial. In these cases, relevant accounts should be defined for each application. As an example, a company can reveal the relevant social media accounts by searching their brand and product names as key words, to analyze the opinion about a specific product, or the change of public opinion about the company itself.

## 5. Conclusion

A social media-based method to retrieve and analyze refugees' issues in Turkey is presented in this paper. We used Twitter as a data source to identify refugees' accounts and how we traced their activities back in time is also explained. A new dataset was gathered; it contained horizontal and vertical dimensions for a multilevel analysis, such as time, location, inference, opinion, sentiment, and politically oriented accounts.

The preliminary study and analysis show accounts that tweet according to events, while other accounts are created to form public opinions about specific political topics. This provides wide potentials for deeper future analysis to understand the conditions of refugees better, and to enhance the public plans for their integration in the hosting societies. To facilitate further analysis, we share our dataset and analysis online [3]. In order to understand the cultural and social value of cultural heritage issues, future work will focus on analysis of cultural elements that the refugees were aware of and concerned about. To represent the collective identity of a community or nation which has lost their country because of war and was forced to leave and take refuge in other countries, by visual means is another future work direction. The interaction of locals with refugees' issues will also be studied and analyzed. A similar approach could be applied on other social networks, such as Facebook, Instagram, etc.

[3] http://ybu.edu.tr/abulbul/custom_page-283-refugee_analysis

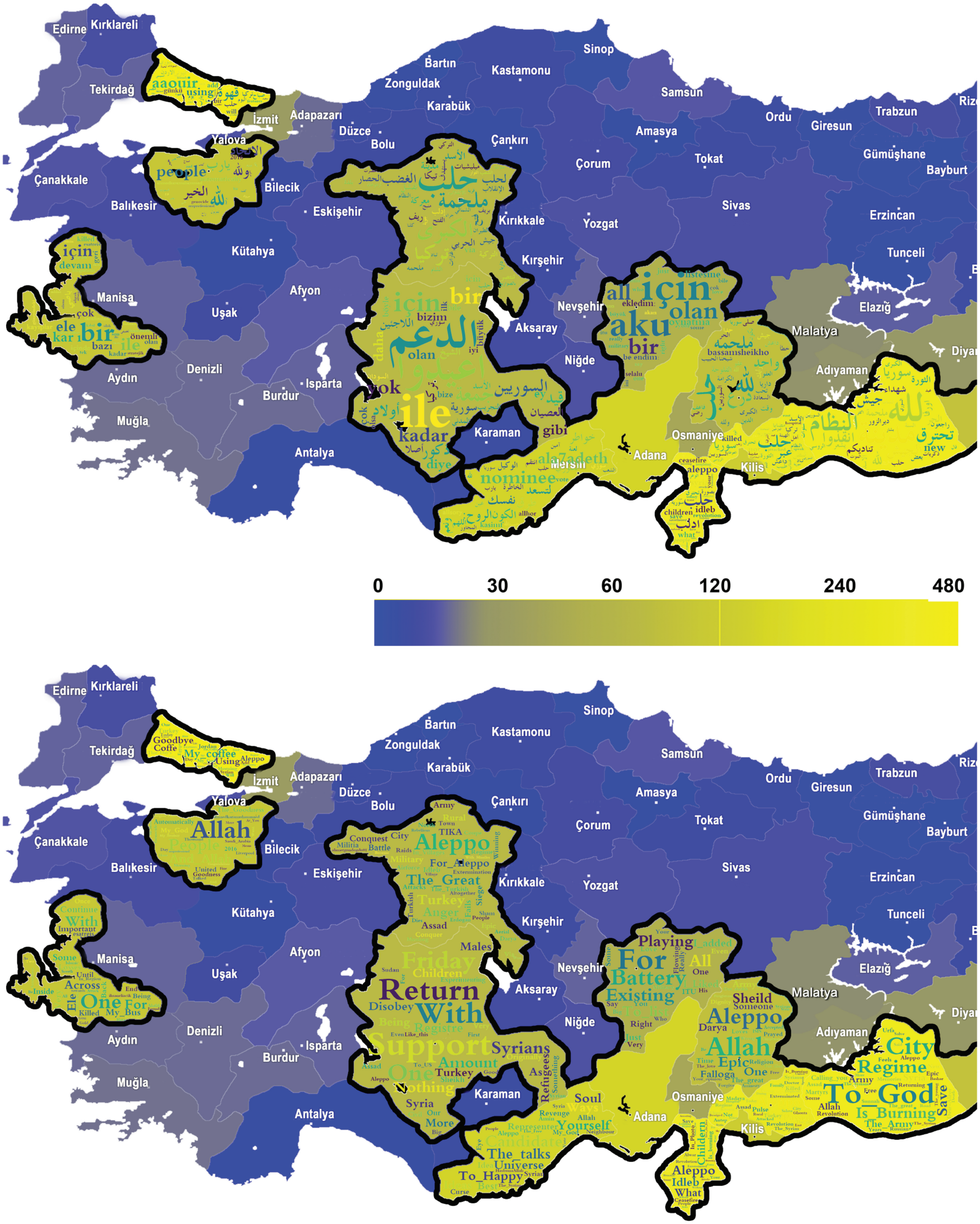


**Figure 6**. Tweets trends in 2016 per city show the difference of popular words in different regions of Turkey. Original, top; English translation, bottom.

The presented research opens the door for a strong involvement and cooperation between different disciplines, and even with its initial results, it shows strong potential for deeper studies in both information retrieval and social sciences.

**Acknowledgment**

The authors would like to thank Mary McIntosh for reviewing the English language of the paper manuscript. This paper is part of the SMART (Social media based analysis of refugees in Turkey) project, conducted in Ankara Yıldırım Beyazıt University.

## References

[1] Stepanova E. The role of information communication technologies in the “Arab Spring”. PONARS Eurasia Policy Memo 2011; 159: 12-16.

[2] Haj Ismail S, Çetin R. SWOT Analysis of refugees’ high education in Turkish universities for construction fields. In: Proceedings of Second International Symposium Syrian Refugees between Reality and Expectation; 20–22 October 2017; Adıyaman, Turkey: Adıyaman University. pp. 93-98.

[3] Jalili M, Perc M. Information cascades in complex networks. Journal of Complex Networks 2017; 5: 665-693.

[4] Helbing D, Brockmann D, Chadefaux T, Donnay K, Blanke U, Woolley-Meza O, Moussaid M, Johansson A, Krause J, Schutte S, Perc M. Saving human lives: what complexity science and information systems can contribute. J Stat Phys 2015; 158: 735-781.

[5] Enli G. Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election. Eur J Commun 2017; 32: 50-61.

[6] Martinčić-Ipšić S, Močibob E, Perc M. Link prediction on Twitter. PloS One 2017; 12: e0181079.

[7] Jalili M, Orouskhani Y, Asgari M, Alipourfard N, Perc M. Link prediction in multiplex online social networks. Roy Soc Open Sci 2017; 4: 160863.

[8] Tumasjan A, Sprenger TO, Sandner PG, Welpe IM. Predicting elections with Twitter: what 140 characters reveal about political sentiment. In: Proceedings of International AAAI Conference on Web and Social Media (ICWSM-10); 23–26 May 2010; Washington, USA: AAAI press. pp. 178-185.

[9] Sang E, Tjong K, Bos J. Predicting the 2011 dutch senate election results with Twitter. In: Proceedings of the Workshop on Semantic Analysis in Social Media, 23–27 April 2012; Avignon, France: ACL. pp. 53-60.

[10] Burnap P, Gibson R, Sloan L, Southern R, Williams M. 140 characters to victory?: using Twitter to predict the UK 2015 General Election. Elect Stud 2016; 41: 230-233.

[11] Mascaro C, Agosto D, Goggins SP. The method to the madness: The 2012 United States presidential election Twitter corpus. In: Proceedings of the 7th International Conference on Social Media & Society: ACM, 11–13 July 2016; London, UK: ACM. pp.15.

[12] Gayo-Avello D. I wanted to predict elections with Twitter and all I got was this lousy paper–a balanced survey on election prediction using Twitter data. arXiv preprint 2012. arXiv:1204.6441.

[13] Kreis R. #refugeesnotwelcome: anti-refugee discourse on Twitter. Discourse & Communication 2017: SAGE Publications Sage UK: London, England; 11: 498-514.

[14] Rettberg JW, Gajjala R. Terrorists or cowards: negative portrayals of male Syrian refugees in social media. Feminist Media Studies 2016; 16: 178-181.

[15] Magdy W, Darwish K, Abokhodair N, Rahimi A, Baldwin T. #ISISisNotIslam or #DeportAllMuslims?: predicting unspoken views. In: Proceedings of the 8th ACM Conference on Web Science (WebSci’16),22–25 May 2016; Hannover, Germany: ACM. pp. 95-106.

[16] Darwish K, Alexandrov D, Nakov P, Mejova Y. Seminar users in the Arabic Twitter sphere. In: Proceedings of International Conference on Social Informatics, 13–15 September 2017. Oxford, UK: Springer. pp. 91-108.

[17] Panetta L. UNHCR Turkey: Fact Sheet. Ankara, Turkey: UNHCR, 2017.

[18] T.C. İçişleri Bakanlığı Göç İdaresi Genel Müdürlüğü. İdare Faaliyet Raporu. Ankara, Turkey: Göç İdaresi Genel Müdürlüğü, 2017 (in Turkish).

[19] Bulbul A, Kaplan C, Ismail SH. Social media based analysis of refugees in Turkey. In: Proceedings of the first International Workshop on Analysis of Broad Dynamic Topics over Social Media: BroDyn, 26 March 2018. Grenoble, France: CEUR-WS. pp.35-40.

[20] Dahab MY, Ibrahim A, Al-Mutawa R. A comparative study on Arabic stemmers. International Journal of Computer Applications 2015; 125:38-47.